DOCUMENT RESUME

ED 195 240 IR 008 977

AUTHOR Hickerson, H. Thomas, Ed.
TITLE SPINDEX Users Conference. Proceedings of a Meeting Held at Cornell University (Ithaca, New York, March 31 and April 1, 1978).
INSTITUTION Cornell Univ., Ithaca, N.Y. Univ. Libraries
PUB DATE 79
NOTE 139p.

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS *Archives: *Indexes: *Information Systems: *Information Utilization: Participation: Records (Forms)
IDENTIFIERS *Information System Users: *Records Management: *SPINDEX System

ABSTRACT

These conference proceedings contain 21 papers presented by users of the Selective Permutation Indexing (SPINDEX) system, designed by the National Archives to produce a variety of computer-generated bibliographic aids for archival holdings. Institutions participating in the SPINDEX users' conference represented government, church, university, disciplinary, and corporate archives and records management programs in the United States and Canada. Papers were presented by users and by SPINDEX representatives in these catagories: (1) current applications in archival control, (2) designing tag (data field) structures, (3) data entry and computer-generated results, (4) SPINDEX III (a system update offering greater sophistication in text and index listings), (5) technical aspects of SPINDEX use, and (6) the future development of SPINDEX. Illustrative materials from the presentations are appended. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED195240

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SPINDEX Users Conference

Proceedings of a Meeting Held at Cornell University

Ithaca, New York, March 31 and April 1, 1978

Edited by

H. Thomas Hickerson

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Thomas Hickerson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Department of Manuscripts and University Archives
Cornell University Libraries, Ithaca, New York 14853

1979

IR008777

Contents

Introduction v

Conference Program vii

Conference Attendees ix

I. Current Applications 1

Evolution of an Archives: From the Covered Wagon to SPINDEX II / Ronald G. Watt

The Genesis and History of the Pacific Northwest Public Power Records Survey / George E. Mariz

Archival Control in the Cornell University Libraries / H. Thomas Hickerson

INCO, Limited Uses of SPINDEX / Richard S. Parmenter

Identifying and Preserving the Documentary Heritage of Washington State / John F. Burns

Archival Uses of SPINDEX II at the South Carolina Archives / Charles H. Lesser

Use of SPINDEX at the National Archives / Stephen Hannestad

The National Historical Publications and Records Commission's SPINDEX Application: A Data Base on Archives and Manuscripts in the United States / Nancy Sahli

II. Designing Tag Structures 36

Tag Structure: NHPRC Guide / Dennis A. Burton

National Data Base Design / Dennis A. Burton

Designing a Compatible Tag Structure for the Statewide Records Survey / David W. Hastings and John F. Burns

III. Data Entry and Computer-Generated Results 55

Data Entry and Output at the South Carolina Archives / Joel A. Shirley

SPINDEX II as a Card Catalog Replacement in the LDS Church Archives / Brent G. Thompson

Data Entry and Computer Output Methods / Richard S. Parmenter

IV. SPINDEX III 67

SPINDEX III / Stephen Hannestad

V. Technical Aspects of SPINDEX Use 78

Cornell Experience with SPINDEX II: Problems and Extensions / Joan M. Winters

Modifications Made by INCO for Converting SPINDEX to RIM/DEX / Richard S. Parmenter

Processing Problems: NHPRC Guide / Dennis A. Burton

Program Additions, Modifications, and Problems at the South Carolina Archives / Joel A. Shirley

SPINDEX II Program Modification / Brent G. Thompson

VI. The Future 91

The Role of the National Archives in the Future Development and Maintenance of SPINDEX / Stephen Hannestad

Appendixes

A. Washington State Historical Records and Archives Project Survey Form 102

B. NHPRC Guide Project Work Sheet 103

C. Exhibits of INCO, Limited Applications 105

D. SPINDEX II at Western Washington University / George E. Mariz 111

E. Guide Project Tag Structure / Brent G. Thompson 117

F. South Carolina Department of Archives and History Exhibits . . 120

G. Designing Tag Structures / Joel A. Shirley 122

Introduction

On the morning of March 31, 1978, forty individuals gathered in the sixth floor conference room of the Herbert F. Johnson Museum of Art on the campus of Cornell University. The sky above Ithaca, New York, was clear and allowed an unobstructed gaze up Cayuga Lake. The group, including archivists, librarians, records managers and programmers, had assembled to discuss the use of SPINDEX (Selective Permutation Indexing), a system designed to produce a variety of computer-generated bibliographic aids for archival holdings. Those attending were welcomed by Donald Cooke, University Vice President for Research, and the introductory session of the SPINDEX Users Conference began.

This conference was not the first time that a meeting had been held to discuss SPINDEX. The first two meetings, which were held at the National Archives in Washington, D.C., March 25, 1968 and May 2, 1969, were planning conferences for those institutions that were cooperating with NARS (National Archives and Records Service) in the initial development of SPINDEX II. More recently, June 11-12, 1973, a conference, also conducted by NARS, was held after the release of an OS (operating system) version of SPINDEX II. Like the earlier meetings, which were unique and denoted significant points in the design and implementation of an archival automation system, the conference of March 31 and April 1, 1978 signified a new stage in the development of archival automation methods.

Unique aspects of the Ithaca conference are that: it was truly a "user" conference, in which almost all of those in attendance represented institutions currently using SPINDEX or in the process of acquiring the system; the twenty-one participating institutions included municipal, state, national, church, university, disciplinary, and corporate archives and records management programs from various areas of the United States and from Canada; the diversity of representation was matched by the diversity of use, including discussion of over forty different applications by the institutions represented; it was conducted by Cornell rather than NARS, indicating an increased willingness among users to take a direct role in the future maintenance and development of SPINDEX, and that, while users will continue to make requests for system support, they will not be completely dependent on NARS for this service; the NARS representative, Stephen Hannestad, provided a clear and positive statement concerning NARS support, brought to Ithaca newly-corrected SPINDEX II program tapes, and presented an introductory description of a significantly enhanced system version, SPINDEX III; and a SPINDEX Users' Network (SUN) was formed and steering group selected. The formation of a user group gave formal recognition to a very important feature of the conference, the existence of a community of users. While the conference was in no way responsible for this development, the commonalty of experience and concern made possible a useful information-sharing process.

The content of the proceedings is diverse, ranging from the general to the specific. The first session is broad in scope, describing the role of automation in various archival repositories. Other sessions concern data base design, data entry methods, finding aid production, technical aspects of system use, and current and future system maintenance and development.

The proceedings are comprehensive, but editorial prerogatives have been exercised. Session and personal introductions have been shortened or excluded. Where papers were submitted beforehand, as was true in most cases, the written version, rather than the verbal presentation, has been published. In Session II, three papers were submitted which were not presented verbally; they have been included in the appendixes. In Session V, five papers were contributed, but there were no formal presentations. Because of the technical nature of the session, these five papers have been reproduced in full, rather than the record of discussion. The record of discussion is included for all other sessions. Some comments have been excluded or rewritten, primarily, in an effort to eliminate some duplication and to convert verbal comments into a written form which is both understandable and, although not stylistically perfect, grammatically acceptable. An effort has been made to preserve a sense of the dynamics of debate and the personal styles of the participants. Although this has not produced the most efficient presentation of facts and opinions, it does reflect the conference environment in a realistic and informative manner.

Conference participants did examine the edited proceedings prior to publication and suggested changes where they felt that the transcription was incorrect or that the record did not reflect the intent of their comments. Such changes were few.

The conference and publication of the proceedings have been supported by a grant from the National Historical Publications and Records Commission. The conducting of the conference and the editing of the proceedings were made possible by the assistance and cooperation of my colleagues in the Department of Manuscripts and University Archives.

H. Thomas Hickerson
June 1979

Conference Program

Thursday, March 30

5:00-7:00 pm — Registration and Open House
Department of Manuscripts and University Archives
101 Olin Library

Friday, March 31*

8:30-9:40 am — Registration and Coffee

9:40-9:50 — Welcome to Cornell
W. Donald Cooke

9:50-10:00 — Conference Introduction
H. Thomas Hickerson

10:00-12:30 — Current Applications
Description and discussion of current applications of the SPINDEX II program package by federal, state, academic, church and corporate archival and records management programs.

Participants:

John F. Burns
Stephen Hannestad
H. Thomas Hickerson
Charles H. Lesser, Chair
George E. Mariz
Richard S. Parmenter
Nancy Sahli
Ronald G. Watt

12:30-2:00 — Lunch

2:00-3:00 — Designing Tag Structures
Discussion of compatible tag structures for records of public and private origin and of the feasibility of institutional, regional or national compatibility.

Participants:

John F. Burns
Dennis A. Burton
David W. Hastings
Brent G. Thompson, Chair

*Friday sessions will be held in the 6th floor conference room of the Herbert F. Johnson Museum of Art.

3:00-4:00	Data Entry and Computer-Generated Results Discussion of data entry methods and forms of finding aid production. Participants: George E. Mariz, Chair Richard S. Parmenter Joel A. Shirley Brent G. Thompson
4:30-5:30	SPINDEX III Description and discussion of further enhancements of the SPINDEX program package and their implementation. Participants: Stephen Hannestad Ronald G. Watt, Chair
7:00-8:00*	Cocktail Hour (cash bar)
8:00-10:00	Conference Dinner
Saturday, April 1**	
8:00-9:00 am	Coffee
9:00-10:30	Technical Aspects of SPINDEX Use Examination of existing program modifications and additions, processing difficulties, and needed changes in system documentation. Participants: Joan M. Winters, Chair Other contributors
11:00-12:00	The Future Discussion of future system maintenance and development. Participants: Stephen Hannestad H. Thomas Hickerson, Chair

*Evening activities will be held in the West Lounge of the Statler Inn.

**Saturday sessions will be held in the West Lounge of the Statler Inn.

Conference Attendees

Brumberg, G. David, Director, New York Historical Resources Center, Cornell University Libraries
Burke, Frank G., Executive Director, National Historical Publications and Records Commission
Burns, John F., Project Administrator, Washington State Historical Records and Archives Project
Burton, Dennis A., Archivist, Guide Project, National Historical Publications and Records Commission
Child, Margaret S., Assistant Director, Research Collections Program, Division of Research Grants, National Endowment for the Humanities
Cooke, W. Donald, Vice President for Research, Cornell University
Dearstyne, Bruce W., Sr. Public Records Analyst, State Archives, New York State Education Department
Dick, Arthur D., Supervisor of Library Services and Records Management, Steel Company of Canada, Limited
Edelman, Hendrik, Assistant Director, Development of Collections, Cornell University Libraries
Colman, Gould P., University Archivist, Cornell University Libraries
Evans, Max J., Chief, Record Processing Section, State Historical Society of Wisconsin
Fagerlund, M. Liisa, Archivist, City of Portland, Oregon
Finch, C. Herbert, Assistant Director, Special Collections, Cornell University Libraries
Hannestad, Stephen, Program Analyst, National Archives and Records Service
Harris, Janie L., Social Sciences Librarian, Cornell University Libraries
Harris, John, Office of Data Systems, General Services Administration
Hastings, David W., Western Regional Supervisor, Washington State Historical Records and Archives Project
Hughes, Cleo A., Director, Program Development and Evaluation, Tennessee State Library and Archives
Jacklin, Kathleen, Archivist, Cornell University Libraries
Jones, James, Office of Data Systems, General Services Administration
Jones, Randall L., Assistant Professor, Department of Modern Languages and Linguistics, Cornell University
Lauman, Frances, Librarian, Cornell University Libraries
Lesser, Charles H., Assistant Director for Archives and History, South Carolina Department of Archives and History
Lloyd, Louise, Records Management Coordinator, Coca-Cola, Atlanta, Georgia
Love, John W., Sr. Programmer, Office of Data Processing, City of Birmingham, Alabama
Lytle, Richard H., Archivist, Smithsonian Institution Archives
Mariz, George E., Associate Professor, Department of History, Western Washington University
Osborne, Jeanne, Systems Manager, Command Records Center, Mississauga, Ontario
Parmenter, Richard S., Records and Information Consultant, INCO, Limited
Sahli, Nancy, Coordinator, Guide Project, National Historical Publications and Records Commission
Saunders, Margaret, Records Analyst, Steel Company of Canada, Limited

Shirley, Joel A., Archivist II, South Carolina Department of Archives and History
Thompson, Brent G., Sr. Cataloger, Historical Department, Church of Jesus Christ of Latter-Day Saints
Treat, William M., Manager, Automated Systems Control Group, Cornell University Libraries
Waggener, Jean B., Director, Archives and Manuscripts, Tennessee State Library and Archives
Wald, Ingeborg, Sr. Manuscript Arranger, Cornell University Libraries
Watt, Ronald G., Assistant Librarian-Archivist, Historical Department, Church of Jesus Christ of Latter-Day Saints
Weart, Spencer R., Director, Center for the History of Physics, American Institute of Physics
Whiting, Marvin Y., Archivist and Curator of Manuscripts, Department of Archives and Manuscripts, Birmingham Public Library, Birmingham, Alabama

Conference Staff

Gustafson, Jane K., Department Secretary, Cornell University Libraries
Herkovic, Andrew, Media Supervisor, Cornell University Libraries
Hickerson, H. Thomas, Associate Archivist, Cornell University Libraries
Pletsch, Ellen J., Catalog Assistant, Cornell University Libraries
Winters, Joan M., Coordinator for User Support, Office of Computer Services, Cornell University

Conference Proceedings

A MEETING OF SPINDEX USERS

Welcome to Cornell: W. Donald Cooke

Conference Introduction: H. Thomas Hickerson

It is very good to see all of you gathered here. I hope that the conference proves to be valuable to all of us. I think that we already represent a community of users, and the question of whether we decide to establish more formal connections is secondary to the fact that we are already joined in an information sharing process. Obviously, in the next twenty-seven hours we are not going to be able to deal with every problem that may arise. Certainly a lot of questions will come up after you leave here, but hopefully, after this meeting, you will know who to contact for certain information, and who is embarking on a project similar to your own. Also, through our own experience and through the distribution of the proceedings, I hope that we will contribute to the general development of effective archival automation methods.

SESSION I

Current Applications

Charles H. Lesser: I think that this morning we are going to get a broad introductory look at some of the big users of SPINDEX. Some of us, including myself, may get totally lost as we get into more technical aspects, but the papers this morning are less technical and provide an overview of what the different agencies are doing and even give us some background on what the agencies are like and why they have gone to SPINDEX.

Ronald G. Watt: "Evolution of an Archives: From the Covered Wagon to SPINDEX II"

Shortly after the establishment of the LDS Church in 1830 the Historian's Office was created by Joseph Smith, the founder and first prophet. He appointed a Church Historian and charged him with the recording and writing of the history of the Church. Because of the short periods these early historians served and because of the persecutions and migrations of the Mormons, very

little history was written until the Mormons settled in Nauvoo, Illinois where Willard Richards was appointed as Church Historian. The process of writing the early history of the Church is described best by Howard Coray, one of the clerks in the office, who remembers that his business was to take the materials furnished by Joseph Smith and combine, arrange, and "amplify not a little, in as good historical style as may be."[1]

In 1846 when the Mormons left Nauvoo, Willard Richards packed all the records in the bottom of his wagons together with his grain and household items and took them to the Great Salt Lake Valley. Ten years later the clerks in the Historian's Office began unpacking the boxes and the process of compiling and writing began all over again.

The main job of the Historian's Office was the compiling of the history which was in a sense a religious duty. The gathering of records was of secondary importance. As early historians wrote, however, they collected documents supporting their histories. Documents which survived the 1846 exodus from Illinois were preserved. In addition, historians sought and received statements and affidavits from men and women who had experienced life in the pre-Utah Church. The Historian's Office, closely connected as it was to the Office of the President of the Church, began to receive documents from the President's Office during this period. Many out-of-date records of the President's Office, the General Tithing Store House, and the Presiding Bishopric were accessioned. The papers of the Presidents of the Church were also accessioned in the Historian's Office. Brigham Young's papers are among the most significant of these. Other types of records slowly began coming into the office, diaries and papers of Mormon leaders and lay members and local congregational records. There was no established policy and records came in haphazardly.

In the late nineteenth and early twentieth centuries Andrew Jenson, an assistant Church Historian, began a process of writing and compiling histories of the local congregations. In doing his research he traveled to each of these congregations. Most of the records he left where he found them, but records in foreign countries he purposely boxed up and shipped to Salt Lake City. Even though Jenson's purpose was to write,he collected more records and manuscripts than ever before.

After Jenson's death a movement began to accession the congregational records. In 1943 the first letter went out to these local units. Most of them dutifully sent their records into headquarters. Now we have a good set of records for the local congregations throughout the world. These records document the activities of the regular member and the activities of each congregation.

Slowly but surely over the years the Church Archives has been taking on the duties and responsibilities of an archives. In 1965 a Church records management program was instituted for Church headquarters. Originally this

[1]Howard Coray, Autobiography, unpublished manuscript, LDS Church Archives.

responsibility was placed under the direction of the Historian's Office, but within the past year it has been moved into another area of the administrative structure. In 1972 the Church Historian's Office became the Historical Department thus representing a broadened function with several divisions including an archives. Several later administrative changes have verified that the archives is a functioning part of the department, and our professionalization over the past seven years reflects this emphasis.

Much work had been done over the years in collecting and preserving records, but professionalism was in abeyance. It has been the professionals who have changed the concepts and brought in the archival theory, so now our staff talks in terms of papers, collections, record groups, series, etc.

It is because of this professional staff that we got ourselves into the use of computers. About six years ago Max Evans and Earl Olson looked into the possibilities of ordering some wooden cabinets for what we call the patriarchal blessing index. They figured that it would take twenty new cabinets which would cost about $1200 each, a total of over $20,000. They budgeted for the cabinets, and then one of the staff in the office who had been a keypunch operator said, "Why don't you keypunch them and put them into the computer?" The Church computer people were actively peddling their wares and were interested in doing something for us. It took some time for the two groups to understand each other, but they finally put together a program, and it worked. They hired part-time staff to input, and when it was finally done the whole thing--wages and everything--cost less than the budgeted $20,000. We now have our patriarchal blessing index on COM microfiche cards, which are updated annually.

We went from that index to other indexes. The biographical sketch index was being kept manually on cards. The same process took place, and we entered the cards in less than a month.

Our involvement with SPINDEX II came as a natural consequence to our relationship we already had with the computer. The whole question was intellectual control of our holdings. We had accomplished this previously through a card catalog. Part of the processing of a collection consisted of filling out a worksheet which required the cataloger to fill in the creator of the collection, the title, added entries, if any, supplied subjects, collation, content note and other pertinent information. Before SPINDEX II our clerical personnel would type a main entry on a simple magnetic tape typewriter. The automatic typewriter would then produce two additional main entry cards and on the average four additional added entry or subject cards. Before 1974 the cards were produced by typing the first and xeroxing additional cards and then manually typing the appropriate heading. The magnetic tape typewriter was a definite improvement from the old system, but was there a better system--a

system that would not require manual filing and would allow for corrections and revisions without pulling all the cards and retyping and refiling them.

When Evans went to the National Archives seminar, he was introduced to SPINDEX II. After he returned he went to our Church computer analysts and asked how much it would cost to develop a program which would give us intellectual control of our holdings. Their estimate was $100,000. Evans decided to work with SPINDEX. The department purchased the program and Evans began experimenting with it. When he left the employment of the Historical Department in January 1977 to go to the Wisconsin State Historical Society, we were ready to begin the converting of our card catalog into a printout form. We thought that the cost of the process would be a trade off. The cataloging time would remain the same. The typist would still have to type or enter the material, but we would not have to worry about the corrections or in filing the cards. What we have discovered is that the time for entering is less than before. We think we have a workable program for much less money.

We have finished with the project of inputting our church records and manuscripts card catalog into the computer and have obtained results. The program itself has given us capabilities that our old card system did not have. We can now generate special lists for donors and languages.

The second project we have for SPINDEX II is that of entering and indexing all of the names of the local congregations called branches, wards, stakes, and missions. The Mormon Church has over 15,000 of these local units. Since we are the repository for their records, we need to have the unit's vital statistics--its organization date, disorganization date, sometimes when it was reorganized, name changes, and the name of its administrative parent organization and dates. Our public services areas need to know this information. We now have all this information in a centralized card file which we are updating continually, but updating is a laborious process of pulling cards, making corrections and refiling. In order to allow each of the public services areas to have their own index and in order to simplify the updating process, we need to computerize it. SPINDEX allows us to do that, and we have begun. It is going to be a gigantic job, but we feel that it will be beneficial and easier to control. For this project we will use both the index and the register portion of SPINDEX.

The department has begun three other applications of SPINDEX. We produced an oral history guide on microfiche that included 283 oral histories, an index, and a project list. In July at the NASARA Convention in St. Louis I learned about International Nickel's use of SPINDEX II in the records management program. We considered that something like this could be applied to our record center which is under the control of the Church Historical Department. One of our general authorities asked for an inventory of his

files in the Record Center, and we used this as an excuse to try it. We have since given him a folder inventory and an index of over 100 linear feet of records. We are also developing a computerized file for our portrait index which is an index to all the identified people in our photograph collections. We already have a card file on this, and we are adding to it weekly. On this project we have test runs back and we perceive no difficulties in our task.

A few other projects are now in the planning stage. Our indexer has started tc index certain Intermountain newspapers. Right now the project is stymied for lack of personnel, especially volunteers to do the indexing. We also plan to index all the reference files throughout the Church headquarters, to institute a cataloging system for the Church museum, and to index the correspondence of our most important manuscript collections. I am not certain though how many of these latter projects will be in the developmental stage by the end of the year.

I feel that the most difficult part of the work is finished. We have made our corrections in the program which make it workable for us, and we have finished the guide project. I see the rest as a downhill process. As long as we have a goal and our personnel working with SPINDEX can remain relatively constant, we can complete most of the projects we envision.

When Max Evans left, I recognized as his successor that SPINDEX had left the experimentation stage and had become an archive project. Brent Thompson, who was one of our senior catalogers of manuscripts, was assigned to take care of the technical part of the development. He needed to work with our processing staff, myself, and the computer programmers. The work from the very beginning has been a collective effort. Thompson has worked full-time on the project for over a year now. He has been responsible for the nuts and bolts. It has been a beneficial experience for us in the LDS Church Archives--a year, for the most part, where we have achieved most of our goals.

George E. Mariz: "The Genesis and History of the Pacific Northwest Public Power Records Survey"

The Pacific Northwest Public Power Records Survey is the brainchild of the Western Washington University Graduate Program in Archival and Records Management Administration. In January, 1976, a number of staff members from the program and Sidney McAlpin, the Archivist of Washington, began to consider the possibility of using Western's graduate trainees as a tool in an off-campus archival and records management project that would be useful to all the parties involved. Initially, they targeted the records of the community colleges, the Public Utility Districts and the local courts as possible objectives. The project's planners eventually settled on the Public Utility Districts for a

number of reasons. First, they were entities with a long history, many dating from the early and middle 1930's. Second, they had a sizable volume of records which were by law in the province of the state archivist, but which still were not under effective control. Individual districts they knew were also encountering records problems and recognized the need for effective records management. Additionally these bodies possessed considerable historical significance and their records deserved to be preserved for posterity as well as for the use of scholars. A brief look at their history reveals the importance of PUD records.

The PUD is a species of the Special District, the most numerous and widespread but least understood of all governmental units in the United States. The Special District is sanctioned by state law but formed by local initiative to perform or provide a vital service for an area which may overlap several, separate governmental units or include only a limited territory within the boundary of a standard governmental unit. The most common is the school district, but water, sewage, library, road, utility and other districts too numerous to mention also are examples of the species. Approximately two-thirds of all governmental districts in the United States are Special Districts, yet they have been little studied until the last ten years.

Special Districts possess several unique characteristics for funding and organizing vital services. First, they have the power to tax their constituents to provide investment or operating capital. However, as they are legal entities created by their constituents, they are also highly accountable to the public. Their board members are either elected or appointed by other elected officials. Thus, unlike private concerns, they are the immediate servants of their customer/constituents and subject to pressures from them. The authorization and formation of a Special District, then, also focuses the issues of public versus private control of essential public services.

The formation of Public Utility Districts in the Pacific Northwest was not only a part of the more general movement for the public control of electrical power, but a movement of immense social and economic significance. Electricity, most especially hydroelectricity, was joined with irrigation and flood control as important matters on the east and west slopes of the Cascades. During the early part of the century the Grangers had pressed for legislation that would combine all three functions, thereby providing them several essential services and, they were certain, cheaper electricity. Legislators in each of the region's states sponsored bills to solve these problems. Private electrical interests, however, consistently opposed such legislation as they were unwilling, and usually unable, to provide the larger solutions for irrigation and flood control.

The issue of the ownership and regulation of electrical generation and distribution facilities arose with the construction of the first electrical generating plant in New York City in 1882. Very rapidly electricity became a necessity rather than a luxury and private companies appeared and obtained licenses from local and state governments unprepared to meet the demands and difficulties the new utility presented. Effective regulation in fact or conception at any level characterized the early history of electrical utilities in the United States. In 1902 the federal government took a tentative first step in the control of the industry with the passage of the Reclamation Act which touched it indirectly. Private utilities opposed any sort of federal intervention and fought public control of the production and distribution of electrical energy; however, with the onset of the Great Depression and the subsequent New Deal the positions of advocates of public or private control became more pronounced.

In Washington and Oregon, which together with the other states of the Pacific Northwest contain approximately half of the nation's hydroelectric potential, the matter was fought with particular bitterness. During the 1920s private utilities opposed any measure to permit the creation of publicly controlled districts. It was not until 1930 that both states passed enabling legislation, but the problems Public Utility Districts faced were still enormous. They lacked the funds necessary to construct generating plants and in a period of severe economic depression had no way to obtain money. Additionally, they confronted continued hostility from private concerns. Finally, they had not yet decided how to confront the intertwined problems of the generation and distribution of electrical energy, irrigation and flood control.

In 1932 many of these difficulties began to resolve themselves with the publication of the so-called '308 Report,' the result of a federal government commission's enquiry into the feasibility of harnessing the Columbia River for multipurpose use. The document recommended construction of a system of dams to provide hydroelectricity and facilitate the economic growth of the region. The Washington State Legislature answered the Report in the following year by creating the Columbia Basin commission which urged the construction of dams at Grand Coulee and North Bonneville. The Legislature further supported the work with a small appropriation which allowed the employment of two hundred and fifty men to do preliminary work at the Coulee site. The size of the work force increased as federal funds augmented the meager Washington State contribution, and by 1934 the Grand Coulee project was employing two thousand five hundred men. The Rivers and Harbors Act of 1935 placed the undertaking in a more secure financial position and allowed the completion of the project. The smaller dam at Bonneville began generating electricity in 1937, and Grand Coulee followed in 1941.

The existence of large, government-owned generating plants brought the heretofore somnolent PUDs to life. By 1937 only four PUDs had been created

in Washington, and only a few more existed in Oregon, but with the completion of the Bonneville Dam the problems of expensive flood control and irrigation were solved, and PUDs could now distribute energy without assuming the burden of generation. Congress passed in the same year the Bonneville Administration Act that provided public control of the generating facilities. The Act insured that the facilities would benefit domestic and rural consumers by reserving a minimum of fifty percent of capacity for them.

Public power distribution, whether by PUD, Cooperative or Municipal System, has continued to grow in the Northwest. Today the state of Washington boasts twenty-two operating and eight other approved PUDs. Oregon has four in the former category and seven in the latter. Additionally, both these states have broadened the base of their public ownership and control by the creation of large numbers of cooperative and municipal systems. Public power agencies distribute almost sixty percent of the electrical energy in Washington and twenty percent in Oregon. Clearly, public power and its most important representative, the PUD, have played a significant role in the history of electrical energy and the overall economic development of the Northwest.

PUDs have also broadened the scope of their operations. Within the two states a number are now members of power generating consortia involved in the construction of new facilities. In this capacity they have become a vital force in the region's economic life by providing abundant, low-cost electrical energy that has lured energy-intensive industries such as metallurgical plants to the area. This in turn has helped to transform the region's economic basis. No longer is the Pacific Northwest an area whose life blood is extractive industries. Primary producers now form an important part of a more balanced and modern order.

The above narrative makes evident much, but by no means all, of the significance of PUD records, as well as the fields for which they have special importance. PUDs keep a number of different kinds of records and not all PUDs have kept exactly the same material. All PUDs maintain records of board meetings, employee records, physical plant records, land title and right of way records, and records of generation, transmission, revenues and customers. Some PUDs have maintained records of the interesting and often involved history of their creation. Their archives, then, are obviously useful to economists and historians. Not so obviously, they are also of value to political scientists, sociologists, demographers and public advocates who may or may not be specialists in one of the above fields.

PUD records, and utility records in general, are of great importance to economists, economic historians and state utility boards. They have provided the information from which rate schedules are compiled and have given state regulatory agencies the primary information necessary to make specific rate decisions and general policy determinations. They have been critically important in the determination of "fair rates." Such records also have been impor-

tant to economists and economic historians, particularly those who use econometric techniques in the study of consumption patterns. Most important, perhaps, PUD records have provided the raw data from which economists are able to make long-range predictions. Various groups, governmental, public and private, have made estimates of future consumption and therefore of future necessary generating capacity. By far the most accurate of these has been the Edison Electrical Institute's (the research arm of a privately owned generating and distributing facility) Forecast in 1959. As of 1971 its cumulative error was only .8%. Significantly, the Forecast used records of consumption by various groups and sub-groups to predict its future needs. These are precisely the kinds of records contained in PUDs, Electric Cooperatives, Municipal Systems and private utilities. Projections by the federal government and other bodies relying solely or even chiefly on the gross generation of power have been far less accurate, and all have required periodic adjustment.

PUD records have also been invaluable to the political scientist. They have been used in studies to illustrate the relationship and problems in state and federal cooperation, the relationship between the special district and the federal government, and cooperation between state and local governments. Recent studies have emphasized the importance of PUDs and other publicly controlled districts in the construction of theoretical models and as prototypes for larger scale public districts that might have future application.

For the historian such records have been a rich source for the study of the region's economic development as well as the conservation and reclamation movements. A large part of the history of the Granger movement still lies buried in PUD records. The organization of PUDs and the election (or choice) of board members is also an important phase in the history of the operation of the New Deal in the Pacific Northwest, though this has, as yet, been little studied. Sources for the study of New Deal support also are available in these records. Of particular interest would be the study of the relationship among local, state and national politicans in electrical energy, conservation, flood control and irrigation policy. PUD records would also be invaluable for biographies of New Deal and later political figures in the Northwest, notably Robert Beck, Guy Meyers and James D. Ross.

For the sociologist PUD records and those of similar institutions have proved a valuable source for the study of attitudes toward public and government control of vital resources. Sociologists have also studied employment patterns in public utilities.

Once they realized the importance and magnitude of the records, the planners entertained the idea of transforming a state-wide project into a regional one. Washington shares the Public Utility District "form" with Oregon, and we decided to contact J.D. Porter, the Archivist of Oregon, to discover if we could expand the area of our work. After a preliminary investigation, he found roughly the same problems and opportunities existed in

Oregon and joined his efforts to ours. Further, we realized that the history of the PUDs was intimately linked to the evolution of Cooperative and Municipal utilities. We therefore decided on a survey project that covered all the region's public power agencies.

Subsequently, we devoted our efforts to liaison with the PUDs. Our first significant contact was with the Snohomish PUD. Members of the project team spent half a day in conference with Margaret Stevenson and Frank Eder, members of the Snohomish PUD staff, discussing our aims and inquiring how we could obtain their cooperation. Of course, we emphasized our interest in producing something useful to them. Our discussions and a subsequent perusal of the Snohomish records indicated we could, indeed, find a common ground for cooperation. The PUD maintained 16,000 cubic feet of records and was increasingly confronted with problems of inadequate storage space and difficulty of retrieval. Snohomish maintained a considerable amount of outdated or duplicate records. The PUD held meterbooks from its inception although any useful or legally necessary information had long ago been extracted. Records of deposits made for connection of service, receipts for payment of bills, general correspondence and many other records had been retained. In addition, the PUD had begun to employ both micrographic and computer techniques to aid them with their problems, but neither had been particularly effective. They filmed many records but continued to maintain both the original document and film. An uncoordinated filing plan rendered computer operations less than ideally effective on all but the most recently created records.

Clearly the PUD required a good deal of help in records management and supported our work when it became apparent to them that we could further the historical part of our work while providing them with assistance.

With a cachet provided by the Snohomish PUD we next approached the Washington and Oregon PUD Directors Associations, describing our work. We found a ready reception there, for it was soon apparent that the Snohomish experience, in greater or lesser degree, was a general one. All PUDs encountered difficulty with the volume and complexity of records they created. This was especially true of the agencies involved in the generation and wholesaling of power. Moreover, even smaller PUDs discovered that new reporting requirements and other recordkeeping duties created similar problems for them, though of lesser magnitude.

Both Associations enthusiastically supported our work.

Armed with a knowledge of our subject and the support of a number of agencies, we then submitted a survey proposal to the National Historical Publications and Records Commission, a branch of the National Archives charged with the preservation of historically important records. The

Commission, fortunately, agreed that our project had merit and funded the work.

Almost immediately, we were faced with problems. First, we had to select and train interns to perform the field work of the survey. Even in doing this we were forced to admit to those we finally chose that we had no intimate knowledge of what lay in store for them. Fortunately, they performed admirably, but more of that later. Additionally, we had committed ourselves to the use of SPINDEX II, the computer program being used by the National Archives and Records Service to produce a uniform national data base. While SPINDEX was to prove a most useful tool and eliminated much of the monotonous and time consuming work, it also posed some problems for us. We were asking it to perform work for which it was not necessarily designed. This is not to say that the work was beyond its capability, but ours was a type of project for which SPINDEX had not necessarily been designed. Our experience with it has on the whole been good, although there are certain changes we shall suggest in a forthcoming piece on our work with SPINDEX.

After we had chosen our interns we put them through an intensive course in archival and records management administration to prepare them for their field work and for professional placement in the future. Simultaneously, we began work on a survey form which would perform the archival and records management tasks we required (Appendix D).

After their academic training each of our interns spent an academic quarter in a cooperating institution where their training was reinforced with practical work. Here they learned firsthand about the management techniques they would have to recommend for PUD records they would encounter. While there, they also performed much of the preliminary work of the survey. They called the PUDs they would be surveying and introduced themselves. They attempted to ascertain the number and location of sites they would have to visit and the approximate volume of records each PUD held. We found wholehearted cooperation at the PUDs, but we also encountered problems, both expected and unexpected. As we already knew, most of the PUDs had no records officers, and records management was often added to the function of an already overworked staff member. In many PUDs, however, no one had exact knowledge of the location of all the organization's records. Except for those with a functioning records program, no PUD had an accurate idea of the volume of records it held. This lack of definite knowledge meant simply that we were often shooting in the dark and the interns frequently had to do a good deal of detective work on-site.

After the completion of formal training, the project administrative staff and the interns convened at the Seattle Federal Archives and Records Center for a final briefing to clear up any problems with the survey form, living arrangements (or lack of them!!), problems to be expected in the field, and other matters that troubled them. We then sent them off to the wilds for the

next three months to complete the field work. Looking back, all of us on the administrative staff owe the interns great respect. Often they survived in the most adverse circumstances. Two of them lived in vans for a good deal of the time in order to conserve scarce travel funds. All kept their expenses to an absolute minimum. In short, their efforts were everything we could have asked and more. They represented themselves and the survey with credit.

Once in the field our interns encountered problems we had not anticipated. While we expected to work at scattered sites, record holding locations were often more numerous than we first believed. Several PUDs had branch offices whose existence was unknown to us. In fact, they were not mentioned in any of the bulletins issued by public power agencies and therefore not recorded in any widely available publication. These exigencies necessitated some shifts in personnel, and we often were forced by circumstances to move interns to work other than that assigned to them in order to complete the survey. Fortunately some kind angels intervened in our behalf. Some of our interns finished their work more quickly than we anticipated and were therefore available to aid others. In addition, the project director relinquished two months of his salary in order that we might use the money to hire two interns during the summer months to finish survey work.

As the survey forms for individual PUDs were completed, they were returned to our office for checking and coding. The data they contained was then keypunched and loaded into the computer on the Western Washington University campus. Printouts of the information provide the basis for the printed material included in the survey. Much of the computerized material does not appear here but has been used to create the other products of the survey, the records schedules and records management training manual.

Now that we have finished the field survey, we can produce a useful document dealing with an important facet of the life of the Pacific Northwest. At the same time, we believe we have met and perhaps solved some of the problems inherent in a field-based records survey, and our successes and failures can perhaps serve as a blueprint for those embarked on similar ventures.

Frequent contact with the PUDs was necessary. Liaison was an important part of the initial phase of our work. We made numerous and frequent contacts with both the PUD directors associations of both states and the individual PUDs as well. Nonetheless, we believe our efforts were still insufficient. Our experience in the field, especially the discovery of so many 'surprise' PUD depositories and sites, means we need to do a more thorough job in this phase of our work. We do not believe that interns undergoing intensive training at cooperating institutions can do more than minimal preparatory work before going into the field. Therefore, we decided in the future the most effective approach would be to employ two advanced students (i.e., two

students holding Masters degrees and whose only course-work requirement was the completion of the archival and records management course itself) as preliminary surveyors. Working in conjunction with the project staff, they phoned various agencies in Phase II of the project and tried to ascertain beforehand which sites presented likely problems. If these could be solved through phone conversations, this was done. Frequently, this was not possible and it was necessary for an intern to make a preliminary visit to the agency to determine all the locations which held records. He then made necessary alterations of schedule, estimates of records holdings and other changes as required. Now, as field work of the second part of our work commences, we have considerably more confidence that we are in control of the volume of work and that we will not encounter the difficulties we faced in Phase I.

Another major problem occurred with the computerization of our data. We had some difficulty with the documentation on the system sent us by the National Archives and Records Service. As stated earlier, we will be publishing a separate work dealing with our experiences with SPINDEX, but the following points deserve some mention in this text. First, NARS documentation does not contain any advice for the unwary. Anything not explicitly noted as within SPINDEX's capabilities is probably not. We urge anyone using the system to refrain from significant experimentation. Most probably it will lead to grief in some form or other. Secondly, the register program format proved to be somewhat inconvenient. To print all the information we required, four separate registers were necessary. This made them awkward to use in conjunction with one another. This was largely the fault of our adaption of the system and can be partially corrected by the creation of a different form. Third, we had some problems with keypunching. This brings up a wide range of matters which should be discussed here. The documentation we received from NARS indicated that only three methods of data input were possible with SPINDEX: punched cards, punched tape and MTST typewritten information. Subsequently, we discovered that other means of input were possible. We suggest that NARS seriously consider updating this section of the SPINDEX User's Guide.[1] At any rate, we decided to use cards as we believed this was the most efficient and economical means available to us. We were forced to rely on the University's Computer Center for keypunching assistance, and initially this worked well. We provided the student keypunchers with elaborate instructions, and they did an accurate, sound job. However, with academic institutions functioning as they do, our regular keypunchers left for vacation when the spring term ended. Summer keypunchers were less well trained and more erratic in their work. While we had only minimal errors in the first twelve thousand cards punched (during the spring term), our rate of error quickly increased, and we had to do a large amount of correcting. In addition, during the summer, we had less time and labor in our office available to verify forms before sending them to be keypunched. If we had spent the time to do this, our rate of error would undoubtedly have been considerably smaller. All this is an elaborate way of saying that good

[1]SPINDEX II: Report and Systems Documentation, National Archives and and Records Service, 1975.

verification procedures and instructions for data processors are a must.

Moreover, we now believe several modes of data input offer distinct advantages over cards. The CPT device, which records data on cassette tapes which then can be viewed with appropriate equipment, provides a means of easy verification of the data before it is produced in printed form. If the project is tied to a card-only computer system, cards can be generated directly from the cassette, thus eliminating any errors involved in key-punching. We are investigating the possibility of entering data directly to a disk pack without the necessity of recording them on some other medium first. If this is possible, we will purchase the necessary equipment since the savings in terms of verification, previewing, and data input time saved are so considerable as to make the equipment cost a small investment.

In many ways, the problems were considerable, but the office staff and interns helped to make it all bearable. We have also learned that a field survey presents many problems and that we have avoided many of the most serious ones.

H. Thomas Hickerson: "Archival Control in the Cornell University Libraries"

As with the group gathered here today, diversity is one of the most important characteristics of the holdings and functions of the Department of Manuscripts and University Archives. Established in 1942 as the Collection of Regional History, the department was created in order to document everyday life in this region. This led to the collection of records of individuals, families and businesses in upstate New York. In 1951 the focus of the department was expanded by the formation of a university archives program. Seeking to document the broad aspects of the university community, personal papers of faculty, students, administrators and alumni, as well as administrative records, have been acquired. The extent of departmental holdings was increased further in the mid-1960's when active collecting efforts were instituted, particularly in the areas of agriculture, urban and regional planning, and political activities. While retaining a regional bias, collecting efforts had become national in scope, and in 1972, the Cornell Program in Oral History was merged with the archival program to form the Department of Manuscripts and University Archives.

While this evolution in organizational structure and collection development policy continues, a similar, but slower evolution has taken place in the methodology for establishing intellectual control of our holdings. Until the mid-1960's, processing and access to "regional history collections" was provided by different staff members from those working with "university archives." Also, different acquisition systems, processing methods and types of intellectual access were used. But in the late 1960's, not only were staff responsibilities integrated, an effort was made to devise a format for analysis and

and description which could be applied to archives and manuscript collections in a consistent manner by manuscript arrangers with varying degrees of experience. It was also hoped that this standardized format could be applied effectively, not only to textual records, but also to photographs, architectural drawings, tape recordings and broadsides.

Of course, it is no coincidence that this standard description format is compatible with computer manipulation, for in 1967, Cornell had joined the National Archives and other institutions in the cooperative development of the system called SPINDEX II. Since then, although the use of a computer-compatible descriptive format has become standard, computerization has progressed slowly, but the concept of computer assistance is firmly entrenched with confident expectation of expanding application.

While this adoption of the computer as a useful tool should not be called a conversion experience, I am impressed by the degree of staff acceptance and confidence. This is illustrated by staff reactions to our newly designed control system for university publications. From its original title, the archival reports and publications project, this system has acquired the catchy acronym, ARP. Although test runs have only been recently completed, I am frequently asked, "Couldn't we put this in ARP?"; "Can I give this to Fran for ARP?"; "Couldn't ARP also give us a chronological listing?" I have to emphasize that just because access to an item is currently difficult, that alone, is not sufficient reason for its inclusion in ARP. So, we will expand computer use with the staff feeling both confident and comfortable.

In my remaining comments I will describe briefly four general types of SPINDEX application. The first two are currently being implemented; the third is currently being designed, and we are only now discussing the fourth. The first is the production of finding aids for particular record groups, specific manuscript collections, or a small group of related collections. Usually seeking access at the folder level, the contents of a folder are described by folder title, chronological span, correspondents, subject, including personal names, institutions, and topics, physical and intellectual record types, and geographical context. We do not attempt to apply all or the same descriptors throughout a collection. Usually this decision is made on a record series basis depending on the nature and access needs of the series. Nor is the same unit for description used throughout a collection. Some series may be more appropriately controlled at the box or item level. Although our descriptive approach is consistent, we retain the flexibility necessary for archival arrangement and description.

A second type of application is the control system for university publications, ARP. Although officially a repository for unpublished documents, our department is also the primary repository for university publications. Upon the establishment of a university archives program in 1951,

an effort was made to acquire published "histories" of various aspects of the University, official publications, and the most common serial publications (yearbooks, directories, etc.). While these titles provide the core, they represent only a small portion of the total number and variety of our holdings. Some items are passed on through the library system; we are on distribution lists for others; and many are received as part of records accessions from university offices.

As our accrual of publications continues, it is apparent that our responsibility for university publications is appropriate. A primary reason for excluding published material from archives, its broad availability, often does not apply because of limited printings and distribution. For some student groups, these publications are the only easily available documentation. Also, publications are frequently useful to researchers using our unpublished holdings.

Unfortunately, acceptance of this role provided us with neither methods nor resources for physical and intellectual control. For primary serial publications, control is maintained with a serial Kardex file. A separate card catalog for pamphlets was established in 1963. This catalogue, applied with consistency, would have provided adequate access, but only 350 titles were ever entered because funds and staff were not available. After a few years, a less expensive process was adopted. Pamphlets were neither catalogued nor listed but were assigned to pamphlet boxes on the basis of their origin within the University. Boxes were arranged in accordance with our general archives control plan. Since access was very poor, the use of this method steadily decreased. During the last few years, publications have been "held" to await future processing.

With the use of the SPINDEX II program package, this year we are implementing a new control system which will substantially improve access. Bibliographic information to be entered and printed in an index will include full title (including successive titles for serials with title changes), author(s), source of publication, publication date(s), whether a serial, a control number (numerically assigned), number of copies in Archives, and, for serials, other library units maintaining copies. Alphabetical access, secondarily sorted by earliest date of publication, will be by title, author(s), proper noun or noun phrases appearing in the title, and source of publication. Textual analysis for subject content is not feasible. When a title is published by Cornell, but no subdivision of the University appears as the source, no source is entered. When proper nouns duplicate other bibliographic data, they will not be entered. Authorship will be designated only when specific individuals are credited. This index will be produced in printout form (11" x 14-1/2") and will be updated regularly. Copies will be available in Archives, major reference areas in the Libraries, and in Interlibrary Lending, a major point of demand. In addition to the alphabetical index, with minor expense, we will produce a chronological

listing and a bibliography. An important aspect of the control system design is the expectation that the majority of university publications will not be integrated into the Libraries' general cataloguing system. Thus, we are designing a system applicable for materials of this type throughout the Libraries.

Currently in the design stage at Cornell is the New York Historical Resources Center. The primary purpose of this program is the development of a statewide historical resources information-sharing system. To this end, a survey of manuscript and archival collections in repositories is being planned. Gathered at the collection/record group level, the data would include information on photographs, architectural records, town views, and other graphic sources, as well as historical manuscripts and records. The major product will be a guide to all manuscript and archival collections in New York State. In addition, each repository which participates will receive an indexed guide to their own holdings and county guides may also be generated out of the data base.

The fourth application is only at the stage of initial discussion and its significance lies in the general approach rather than application design. It also illustrates an important point relative to the developing debate over a national bibliographic system for archives and manuscripts.

It has been suggested by library staff outside of our department that some groups of publications should be analysed and described as archival collections rather than catalogued on an individual item basis. Not only is economic feasibility sited, but it has been suggested that this type of access is more appropriate for some types of publications. As with archival materials some publications are only valuable in context with publications with which they share a common creation. This is the case with some series of governmental publications, some periodical collections, and some groups of microform publications. Another case sited is a group of Vietnamese novels, individually of little value, but as a collection, they provide valuable documentation. In this case individual item cataloguing is inappropriate and the high cost of foreign-language cataloguing is unjustified. Also, as with archives, Hendrik Edelman suggests in a paper presented at last summer's meeting of the American Library Association, publication source may be the most important aspect of bibliographic control.[1]

I am sure that, for some archivists here, this discussion is somewhat paradoxical, for we have all been told of the destruction wrought by librarians treating the documents in a collection as discrete items. While we have moved toward a state of peaceful coexistence with library methodology in recent years, I find these developments quite interesting. It should also enlighten our thinking concerning a national bibliographic system, for this discussion is going on in our library at a time when re-

[1]Hendrik Edelman, "Selection Methodology," Paper Presented at the American Library Association RTSD-RS Collection Development Preconference, 1977.

placement of the card catalogue by an on-line catalogue is being seriously investigated. Yet, it seems apparent that the same type of access and control is not appropriate for all publications. It is also apparent that as with archives and manuscripts, neither the nature nor the cost of on-line access is always justified. Hopefully, we will also approach archival automation in a critical and creative manner.

Richard S. Parmenter: "INCO Limited Uses of SPINDEX"

In 1970, the Corporate, Information Communication and Records Management group was searching for someone who was converting a permutative index producing computer program known as KWIC for use on the then new IBM 360 series computer. Inco discovered SPINDEX II at the National Archives and acquired an early (DOS) version in April of 1971. In the seven years since, we have established some thirty applications--each with two to twenty separate lists or indexes.

Here is a listing to see the diversity of this SPINDEX computer program:

RIM/DEX Applications in Current Operation or in Testing

Books
Government Specifications
Company Marketing Publications
Legal Opinions
Industrial Market Classifications
Computer Programs
Manuals
Forms and Reports
File Folders
Legal Documents
Press Releases
Proposed File Classification Plan
Record Center
Photographs
Engineering Drawings
Engineering Projects
Capital Appropriations
Offset Printing Plates
Records Inventory
Alternative Information Sources
Slides and Presentations
Artwork
Legal Requirements
Approved Records Retention Policy

At this time, some seven years later, we are very happy with the results and predict the uses will grow even more rapidly over the next few years. It is by far the most cost effective of the many programs we have investigated and the few, which we have placed into operation. Clearly, this set of programs known as SPINDEX II is the most practical starting place for any organization interested in finding and controlling its paperwork. If an organization could have only one program, this program would carry it further toward its objectives than any other computer program of which we are currently aware.

Inco's objectives are geared to the Information Resource Management concept, which is very large in scope and extremely diverse in content.

The Information Resource of INCO Limited, for instance, includes all correspondence, all reports, all forms, all directives, and all printed literature, all visuals, all audios and all machine language media. It includes all of these items over a life cycle from creation or receipt, thru processing and distribution, thru organization and retrieval, and finally thru retention and disposal.

For Inco, this is over one thousand million pieces of paper, film and magnetic media distributed in over 100 geographic locations in over twenty-eight countries around the world. This would include archival material which is less than three percent of the total.

We want to know what we have, and where it is located. No small task--probably accomplished by few, if any, organizations, yet, potentially possible with the aid of SPINDEX. Even more important is continuous, daily recovery of miniscule segments of this Information Resource to help our 50 thousand employees perform their daily work. Again, SPINDEX can aid greatly. In addition, the Information Resource acquires new items at a rate of 50 to 150 million, or five to fifteen percent of the total resource, each year and hopefully retires a similar amount. SPINDEX has proven invaluable in allowing us to see that what we are doing is what we are supposed to be doing in the management of this valuable resource.

These are not new goals. What is new is the aid of computers and particularly SPINDEX to help in achieving these objectives. Up to now, it seemed like a losing battle to the paperwork explosion. Now it seems possible to wage a winning battle.

Our experiences with these many applications and reports has prompted us to set up a simple classification of basic, intermediate or advance applications. All applications from basic to advanced run exactly equally on the computer under the SPINDEX programs. The differences from basic to advanced applications occurs in the system analysis, the preplanning and the experiential skills required of the individual setting up the application

procedures, and the formats of the lists and indices to be produced.

PHOTOGRAPHS

The photograph application is an example of a basic application. The objectives of a basic application are:

1) What do we have and where is it?
2) How to find what we have?

There are two computer products, one a listing in chronological order of the photographs showing the identification number of negatives and prints, people names, occasion or situation descriptions, date, color vs. black and white, and size. The second is an alphabetized keyword index with the keywords being generated by the computer from the people's names and occasion or situation description. The entries under the keywords show the same information as the other listing.

DOCUMENTS

The document application is an example of an intermediate application.

The objectives of the intermediate application include the basic application objectives, namely:

1) What do we have and where is it? (Table of Contents)
2) How to find what we have? (Keyword Index)

PLUS

3) When should it be reviewed?

OR

When is it terminated? (Control Lists)

OR

Who is the legal reviewer?

FORMS PRINTING SPECIFICATION

The forms printing specification is an example of an advanced application.

The objectives are the same as the basic or intermediate applications. However, the items in the collection have so many descriptors, it is only practical to list them vertically rather than horizontally. Therefore, the control numbers are assigned out in blocks so that each item is represented by two or more control numbers. The number of descriptors may vary from item to item.

This represents a major step forward in controlling multiple collections in multiple locations that pertain to a common item, such as one employee or one piece of equipment.

Unless you have reason to be familiar with many of the computer programs within an organization, it is unlikely that you will appreciate how different from the usual, the SPINDEX programs are. It is important to recognize and understand those features which make it so unusual.

UNCOMMON FEATURES

1) Computer-generated and/or assigned index terms.
2) Multi-purpose, canned program.
3) Add applications or reports yourself.
4) Change report formats yourself.
5) No training of computer people.
6) People oriented and optimized.
7) Accommodate any existing codes.
8) Codes and/or words.
9) Variable number of fields, length of fields and number of entries per field.
10) Uses full typewriter keyboard.

Whether your organization has no information retrieval type computer programs, or whether it has many sophisticated programs, SPINDEX will help. It would be hard to get a better value for your money.

John F. Burns: "Identifying and Preserving the Documentary Heritage of Washington State"

In July 1975, when Governor Daniel J. Evans was invited by the National Historical Publications and Records Commission to appoint a State Historical Records Advisory Board, there was little knowledge of the extent and condition of historical records outside archival custody in the state of Washington, nor much opportunity to gain such knowledge. The three major repositories, State Archives, the University of Washington and Washington State University, were far too under-staffed and under-financed to even envision accomplishing such a task as a statewide survey. Yet, not three years later, such an investigation was well underway. This essay will trace the development of what is now called the Washington State Historical Records and Archives Project, and outline its goals and its experiences to date.

The State Historical Records Advisory Board, chaired by State Archivist Sidney F. McAlpin, first met on March 30, 1976. Richard Berner of the University of Washington opened discussion of the development of a state plan to identify priorities within the state, and suggested that archival

training, followed by a survey and accessioning project, might be the proper approach. Establishment of priorities was further discussed at subsequent meetings, with archival training workshops followed by a statewide survey project heading the list.

Berner and McAlpin authored the first draft of the State Plan in April 1976, proposing that public and private records be surveyed throughout the state, with a concentration on local public agencies, businesses and subelite organizations, all of which had for the most part been ignored by existing repositories. Sidney McAlpin at this point began to assess the magnitude of the task at hand. Noting that there were 1,717 various and sundry governmental jurisdictions in the state, ranging from massive county governments to six mosquito control districts, McAlpin wryly observed that "it is important to have a grasp of how mammoth the dinosaur is before we start kicking it." Further analysis of priorities was clearly in order.

Through the summer of 1976 and into the fall the Board continued its discussions and made several revisions to the initial draft. The dinosaur was indeed so "mammoth" that NHPRC recommended that the Board attempt to survey only half the state in the first year. The Board decided to work primarily in areas where a college or university could provide central support, to reduce the number of educational workshops from an original fifteen to six, and to delay guide production and accessioning plans until a later time. The final draft of the plan was submitted to the Commission in the winter of 1976, with implementation targeted for Spring of 1977.

The Board continued to work through the winter on preparation for implementation should funding be granted. Nevertheless, when formal notification of funding was received in March 1977, much remained to be done. In complying with prevailing standards for equal opportunity and affirmative action the Board faced considerable delay in implementation as, realistically, little work could transpire prior to the appointment of an administrator and survey team leaders for the project. The hiring process was completed in June and it was only then that the project was actually inaugurated, initially in a storm of manic activity on the part of supervisory personnel.

Predictably, several problems were immediately encountered. Attendance at the educational workshops was often low, as prospective attendees were frequently on vacation, filling in for vacation-bound colleagues, unable to find travel funds or inundated with other meetings. Professional records custodians were inclined to attend only sessions that would solve their most pressing problems. One county officer, for instance, traveled a lengthy distance to hear a one hour presentation on microfilm. It was later learned that a few days previously he had been set upon by a rather insistent microfilm vendor. A fair number of attendees consisted of merely interested (or disinterested) citizens, with a small cluster of vocal crackpots to boot. One lady, for instance, requested that we intervene to save her records from the "mad arsonist" who, at periodic intervals, would infiltrate her attic

with a blowtorch and singe the edges of her records. Another insisted that I reveal the "true" connection between the project, the Commission and the CIA!

As a result of experiences with the initial workshops, the Board considered scaling these down in terms of subject matter and duration. Accordingly, the final week-long workshop was divided into two two-day affairs, and was much more productive as a result. Another aim of the workshops had been to identify volunteers to perform survey work. This was largely a hopeless proposition as, with rare exception, those people in a position to do volunteer work were severely limited in their inherent abilities to capably handle such work. The workshops did, however, attract some paid personnel to the project, were useful in training these personnel, and were of considerable value to those records custodians who were in attendance.

Due to the time spent on workshops, survey work was not completely underway in all regions until mid-September. Even then, progress tended to come slowly. Inevitably, confusion reigns triumphant at the start of any new project. Many hours were spent revising survey form instructions, resolving points of dispute, greasing the creaky wheels of the administrative chain of command, and instructing surveyors. It was found that resolution of these problems could only come through experience, as insufficient data existed to make thorough plans. Considerable personnel turnover also began to occur as the type of intelligent and professional people who make good surveyors would not work at $3.75 an hour if a better opportunity presented itself. The use of SPINDEX-based forms requires a high level of precision and accuracy, and the survey time frame was so short that little extensive training could be given. For that reason people with historical or archival backgrounds were sought, even though they might, and frequently did, leave the project for other work, or were encumbered by other part-time jobs that competed for their time. At times the project staff believed they should have taken the advice of Professor Herman Deutsch, who directed the WPA survey in Washington in the 1930's. He claimed that a few unemployed "ladies of the evening" were among his best surveyors.

Despite the problems encountered, the pace of survey work accelerated rapidly, and a core of highly talented people remained committed to the project. By March 1978, close to 40% of the state had been surveyed, with data recorded for over 540 institutions, 3800 record groups or collections, and 6800 series, the latter almost exclusively in the public sector, as few private collections required series level inventory, but most record groups did. Priorities for survey continued to be refined, in the end focusing on county and city governments, selected special districts, major repositories and libraries, historical groups and a selective sampling of private organizations.

At the administrative level, considerable effort was devoted to reviewing forms and compiling survey-based data to track time and costs. All

forms were initially checked for accuracy by a survey leader, and were rechecked for completeness of format by the project secretary. Survey leaders were also required to file monthly reports which provided not only a narrative of the month's work, but also a statistical breakdown of survey time, and thus related costs, for the various sectors--governmental units of various types, private repositories, businesses, churches, community organizations and so forth. At the conclusion of the project this data may be useful to other groups contemplating such an undertaking.

Preparation is now well underway to extend the survey to the entire state, with completion of the data collection phase scheduled for early 1979. The survey forms have recently undergone their third, and hopefully final, revision. Some of the administrative paperwork has been streamlined, and the supervisory structure has been revamped to reduce the number of supervisors and increase the number of paid hours available for surveying. Grades of surveyors have been established to reward the best personnel with promotion and corresponding salary increases. This perhaps will reduce the number of untimely resignations that have been a source of frustration, and enhance the overall quality of the survey work.

As before, the survey will begin in each region with workshops for public officials and local historical groups, though with a more concentrated approach than last year. These will often be retitled "public meetings" for reasons cited by Larry Stark in his final survey leader report, in which he noted that "the very term 'workshop' may have been a poor choice" for his sector of the state, where the word "connotes bureaucracy and intrusion . . . and may be generally offensive to the common social ethic." Half-day intensive sessions for each sector will be held. These will provide information on state law and records management for public officials, offer a brief introduction to archives for private groups, and explain the survey to both audiences. In fact, nine sessions for public officials have already been held, with over four hundred people in attendance. These results are far more promising than those experienced in the previous year.

Editorial work on the raw survey data will begin shortly and will proceed concurrent with completion of the survey. Significant effort in this area will be necessary, as it is nearly impossible to achieve a totally acceptable degree of quality control when utilizing over fifty personnel of diverse background and experience to collect the data. The work of volunteers is especially imprecise, and some forms will likely be discarded. In certain cases the zeal of volunteers carried them far beyond the bounds of at least this type of survey. Probably the most peculiar example of this to date was the submission of a form entitled "Bones Found on Judson Street." Even more distressing though is the fact that the surveyor then failed to provide an accurate description of the bones or their species of origin.

Initially, survey forms for one three-county area will be edited, and a small test guide of about 1500 entries will be produced. This region was selected because the survey there is complete, and because the area is sufficiently representative of all types of institutions and records that are likely to be encountered. On the basis of the experience gained in compiling this initial publication, and presumably in surmounting the difficulties that will doubtless engage the staff's attention, editorial preparation will continue and data entry will begin for the rest of the state. Even with the "bugs" squashed and the operation running smoothly, this is expected to be an extraordinarily time-consuming endeavor, extending into 1980. Rousing the dinosaur is one thing; taming the beast is quite another.

By some point in 1980 the indexed guide to records in the entire state should be completed; it is likely to be a publication so immense that it will appear only in microfiche form. In addition, more specific guides will probably be produced in hardcopy for specific regions of the state, subject areas, repositories, and types of institutions or records. Plans to devise a program by which the guides will be periodically updated are also being made. Updates will reflect new accessions at repositories throughout the state, including not only the major ones, but minor organizations at the local level as well. The intent is to induce their cooperation by offering them technical advice and on-site consultation in return for their agreement to participate in an archival reporting network, the full details of which have yet to be developed by the Advisory Board.

Eventually, the Board hopes to realize a dream that it has been fortunate enough to have been given the support to undertake. Richard Berner noted in a memo to his colleagues on the Board in April 1976: "I hope all of you share with me the view that this is the first real opportunity for interinstitutional cooperation and community involvement in the archival field since the Historical Records Survey (of the 1930's)." And the community response has been most encouraging. To quote Ms. Kathryn H. Martin of the King County Historical Association: "What the survey has brought out is not only the need for locating and preserving archival records but the need for their proper organization . . .. It is surprising for even a comparatively small society like ours how many requests we get for local historical information and the potential for genuine research that we hold." She goes on to say that the existance of the state plan "was the most encouraging news I'd heard for those of us struggling to preserve valuable historical records."

Of course, much more is involved in this than merely advancing the state of archival work. The principal goal of this effort is and must be service to researchers, professional and amateur, of whatever interest,

throughout the state and nation. In the recent past many commentators have postulated that the United States is floundering, unable to solve its pressing problems by virtue of its inability to clearly relate the past to the present. While this paper is not about to address that specific assertion I think it is important to note that it is a renaissance of interest in state and local history and the possibilities inherent in it that makes this project so consequential. As Clifford Lord observes in Teaching History with Community Resources, with "localized history, not only do you find local, tangible illustrations of the national generalization, but you find other developments which are quite the opposite. From this comes, obviously, a healthy skepticism about the glib generalization and a new appreciation of the diversity of the American scene." Localized history studies might also effect a better appreciation of the continuities of human experience, and thus provide a balance to the careless, present-minded generalities and clichés which have too long characterized American life. If the project can, in some small way, assist movement in the direction that Lord finds so meritorious, it will in sum have been worthwhile.

Charles H. Lesser: "Archival Uses of SPINDEX II at the South Carolina Archives"

Archivists and manuscript curators are notorious for each marching to the beat of their own drummer. Older institutions frequently inherit a hodgepodge of finding aids, no two of them alike; and the archival world is far behind our colleagues in the library profession in developing consistent access tools and shared data bases. One of the purposes of this conference is to explore the possibilities for compatibility of computerized archival data banks. At the South Carolina Archives we have developed a coherent, integrated finding aid system about which we are none too modest. Nonetheless, at the outset I must report that we too have been marching to our own drummer.

South Carolina is fortunate in possessing one of the richest sets of public documents of any of the original 13 states, but these records are still under sadly inadequate intellectual control. Perhaps, a bit of history will put our current efforts in perspective. The institutional origins of the South Carolina Department of Archives and History can be traced back at least to a "Public Records Commission" created in 1891 to secure transcripts of South Carolina materials in the Public Record Office in London; however, an adequate building was not obtained until 1960 and most of our 12,000 cubic feet of governmental records were not gathered from attics, basements, courthouses, offices, and warehouses until very recent decades.

To make matters worse, most of the loose papers that did come into archival custody at an early date were processed during the first part of this century without any regard for provenance and original order. One of our current major projects involves restoring and reprocessing a massive group of our most important antebellum documents that were filed in a subject scheme in acid manilla folders and ordinary filing cabinets. Unfortunately, we still do not have a sufficiently large staff to adequately process, or in some cases reprocess, our tremendous backlog of public records.

Several factors make life easier, however, as we attempt to deal with approximately twenty-four million sheets of paper dating from 1671 to the very recent past. For one thing, we do not collect private papers and can therefore focus our energies on the state and local governmental records that are our reason for being. For another, South Carolina is a small state with only 46 counties and a total population of only 2,848,000. As a proprietary and later royal colony, South Carolina did not develop effective county government and recorded all legal instruments in Charleston. After the American Revolution the keeping of public records did decentralize among a proliferation of subdivisions of local government, but the state's small size makes less visionary our goal of assembling in the Archives Building in Columbia, either in original form or on microfilm, virtually all historically valuable governmental records, local as well as state.

Faced with a mountain of inadequately accessible public records, we have developed a finding aid system with three levels. The first and most cursory of those levels is our published Temporary Summary Guide. A simple list of the record series in our holdings arranged according to provenance in record groups and subgroups and noting dates and quantity, this guide is designed to be frequently superseded by new editions as we make progress in gaining control of our holdings. The second edition was published late in 1976 and a third edition is now on the drawing board. If we should be so fortunate as to ever attain our goals for the second level of our finding aid scheme, a computerized descriptive guide to our holdings, the Temporary Summary Guide might be superseded altogether.

After such a lengthy introduction, you may be relieved to know that I have finally gotten to a discussion of the first of our two uses of the SPINDEX programs. This use, which is the second level of our finding aid scheme and which we call our "Descriptive Guide," is in essence simply a series level prose descriptive finding aid similar in content to the published inventories of the National Archives and Records Service. The detailed guidelines that we have drawn up for our archivists specify that for the 900 tag they are to write up to five paragraphs of no more than 564 characters each. These paragraphs, which are in the form of clear standard sentences with only limited, prescribed abbreviations, contain a carefully

researched description of the series and must include such data as provenance; statutory citations for the creation and changes in the series; description of the content of the series; mention of unusual items of special note; the form and arrangement of the series; description of existing indexes and other finding aids; citations for microfilm, transcripts, and publication of the series; and indication of any restrictions on access. At the South Carolina Archives we feel strongly that we will not have fulfilled our obligation to the thousands of record series in our custody until we have painstakingly prepared these descriptions for each series. Unfortunately, I must report that at present we have only about 200 draft series descriptions written, although it is true that these drafts are for some of our most important and complicated records. Both the detailed instructions for preparing the input for the guide and the necessary program modifications have been ready for some time, but none of the entries for this guide have yet been encoded.

With this audience I need not dwell on the advantages of using SPINDEX for such a guide. Perhaps it is worth mentioning here, however, that a significant portion of the record series with which we deal are still active in the agencies that created them; ease in updating entries as additions to series go through the records management cycle and come into archival custody is therefore a significant criterion for us. The assistance of the computer in processing the indexing terms that we select for each of these series descriptions is also an undeniable advantage. Still it is true that such a guide would be feasible using conventional, noncomputerized procedures. Since we have access to both the hardware and software for a computerized system, we are taking advantage of that access. We have set up our output to print in archival order first a general description of the record group which includes discussion of what subgroups, if any, are in that record group; then a description of the subgroup that includes treatment of how the series are arranged within that subgroup; and finally, the series descriptions. At the end an index will appear which will print out series titles and dates as well as entry numbers with each index reference. You have all been given a copy of an old mock-up of how our descriptive guide will look. We are going to produce our output for the guide in the form of 16 mm computer output microfilm for the forseeable future. When we have substantially completed describing our holdings, we might hook up our tapes with a program for electronic photo-composition and issue a letterpress guide in place of our Temporary Summary Guide.

Much of modern conventional archival wisdom cautions against attempting to gain item level control. Document or case file level indexes and calendars, so this argument goes, are prohibitively expensive and should not be attempted. It is precisely in bucking this conventional wisdom that the South Carolina Archives is making its most radical and most successful use of the SPINDEX II programs. We are spending more than $60,000 a year

on the third level of our finding aid system, an item level computerized data base that we usually call our "Document Index." Since July 1973 we have encoded indexing information for over 100,000 archival documents or case files. In the nearly five years that our project has been operational we have sufficiently established our procedures and have so significantly improved our efficiency that at our present rate we will add 50,000 documents to our index data base each year. In the very near future we will have available in our Search Room a completed, consolidated computer output microfilm alphabetical index to the 70,800 documents in seven record series: Judgement Rolls of the South Carolina Court of Common Pleas, 1703-1790; microfilm of the John McCrady Plat Collection, 1690-1929; the Conveyances recorded by South Carolina's colonial Public Register, 1719-1775; the Renunciations of Dower of the South Carolina Court of Common Pleas, 1726-1775, and of the Charleston District and County Court of Common Pleas, 1776-1887; our Accounts Audited of Claims Growing Out of the American Revolution, 1776-1860; and a series of land grants, 1694-1776. An additional temporary partial index to the portion that we have completed of our antebellum Petitions to the General Assembly and Reports of Legislative Committees will swell the number of index references available to researchers to nearly a third of a million. After each and every one of those third of a million index references will appear not only control numbers that provide access to the documents but also a concise abstract of the document.

A reputable beginning, you may say; we might immodestly agree. Extravagant, you may say; we think not. The reasons why we think not, in addition to the simple fact that we set very high goals for access to our holdings, lie in the nature of our holdings and in the nature of our reference demand. As the governmental archive for the state of South Carolina, we are heirs to a spectacularly rich body of documentation about virtually all people who in three centuries have lived in the state, from Nero, a slave carpenter, to James Glenn, a colonial governor, and about almost every conceivable aspect of the state's history. Many public records by their very function when they are active must be easily retrievable; a probate judge, for example, must be able to quickly locate the inventory of John Jones' estate. With our some twenty-four million sheets of paper we inherit literally thousands of indexes.

Much of the user demand from the persons, academic historians and genealogists alike, who make more than 10,000 research visits to our Search Room each year is also for the specific item, not the whole series. While scholars may well expect to have to read the personal papers of a leading literary figure from beginning to end and in many cases will want to anyway, the case is usually otherwise with the massive series of governmental documents in our custody. Much of the best of recent scholarship in American history reflects new emphasis on public records; the host of community studies, the exciting things that prosopography is revealing, and the urge

to do history from "the bottom up" come immediately to mind. Our thousands of disparate indexes, most of which provide access only through personal name, and hosts of critically important series with no indexes at all are quite simply not adequate in our view. And so we go valiantly forward, hopefully not resembling Don Quixote, towards our goal of a massive consolidated index at the item or case file level to the most important of our early record series. As well we are developing quite a variety of other specialized indexes generated from the same data base. And what is more, we at least talk about the day when access to that data base might be on line.

Existing indexes once they are encoded in many cases can do much of our work for us. While arranged only by personal name, many of our existing indexes also contain date, geographic location, quantity, and a variety of other data which, when converted to our SPINDEX format, can become index references. To cite one example, one of the most exciting things that we are just now gaining for the first time is limited geographical access to our remarkably complete colonial land grants, access that was buried there in the 19th century index all along.

In both of the SPINDEX applications that I have been describing we have been careful in designing our system to allow for it to become a statewide data base that could include the manuscript collections in all of the state's repositories as well as the governmental records with which we work. A statewide network, however, is the upper limit for the system since we are utilizing a unique tag structure and have devoted a great deal of effort to developing a highly specific subject thesaurus, large parts of which would not be applicable to the sources for the history of other states. Two handouts that you have and another paper which you will hear later, all by Joel Shirley, will enlighten you on some of the more technical aspects of our uses of SPINDEX. I need only spend a minute or two on our subject thesaurus at this point.

Despite what I have been saying compatibility has always been an important consideration as we struggled to develop our system at the South Carolina Archives. Naturally then, as we began to think about series for which we wanted to include topical as well as personal name, geographic, and other forms of indexing, we turned to the Library of Congress subject scheme used almost universally for books. It did not take long, however, to confirm what we have since learned has universally been the case elsewhere: this system is not satisfactory for manuscripts. As time went on it became clear that no existing thesaurus would meet the needs of our project. Relying especially on Dagobert Soergel's book Indexing Languages and Thesauri: Construction and Maintenance but also using whatever else we could lay our hands on, including the L. C. book subject headings, we have developed and continue to expand and refine our own thesaurus for subjects. And, of course, we also use authority files for geographic locations and

legal types of documents. I cannot emphasize strongly enough how important, at least for the sort of massive data base that is actually coming into being in South Carolina, the required use of thesauri is. Indeed, one of the chief virtues of going to the computer for our operation, I believe, is that it forces us into a rigidly consistent system of processing our holdings.

After nearly five years experience and concentrated effort our overview of our SPINDEX operation is one of great excitement. Surely at this conference we must at least consider reformatting our descriptive guide to make it consistent in tag structure with the new national guide data base being developed by the National Historical Publications and Records Commission. Since we have not yet begun encoding for that level of our finding aid system we may not yet have passed the point of no return. Surely, too, we must redouble our efforts to speed our progress on that level of our finding aid system, the level to which conventional wisdom points but which we feel we must far exceed.

As to our document index, perhaps it does not matter as much that compatibility may be largely limited to one state. I can assure you, however, that it is a most exciting thing as the computer whirs and grinds and is already bringing into being what is truly a revolutionary improvement in access to the remarkable body of records in our custody.

Stephen Hannestad: "Use of SPINDEX at the National Archives"

The National Archives is using SPINDEX in its most traditional intent, strictly to produce indexed guides to our holdings. The two current projects at the main archives, which are using SPINDEX, are the Guides to German Records that were microfilmed at Alexandria, Virginia and the Index to the Papers of the Continental Congress. The Index to the Papers of the Continental Congress was a detailed item-by-item description that produces both a subject index of approximately four volumes, roughly 8,000 pages all told, and also, for the first time, a complete chronological list of all of the Papers of the Continental Congress. Hopefully, we will be getting this project into the publication stage later this year, and it should be on the market in time for some Bicentennial-related event, perhaps the Treaty of Paris.

The Continental Congress Papers project has been challenging. It has required over six years of work; many archivists have worked on the input. It shows how SPINDEX can be used in an extremely important record holding to give very detailed indexing, and then we turn around, and with another project at the National Archives, actually in the Commission, we use the same system for a broad national guide. This shows the realm of capabilities within strictly traditional archives description using the SPINDEX system.

The Guide to German Records is a case where SPINDEX was asked to continue a project which had been begun manually. Under manual production, we had been able to produce, with varying frequency, guides without indexes. Since 1974, we have produced ten new guides in this series, and we are now working on the production of a consolidated index to these last ten guides. At the National Archives we are now considering SPINDEX application for a variety of other projects, including the indexing of pension records and some of our ordinance records.

Nancy Sahli: "The National Historical Publications and Records Commission's SPINDEX Application: A Data Base on Archives and Manuscripts in the United States"

The National Historical Publications and Records Commission's current SPINDEX application is the preparation of a data base on archives and manuscript sources in the United States, and publications, such as the Directory of Archives and Manuscript Repositories, which will be published later this year, derived from this data base. For the Directory, we have relied on the SPINDEX program package for two basic operations: manipulation and arrangement of the data, and, by using interface programs, production of the printed product, a sample of which is being passed around the room. Of course, the SPINDEX program itself does not include the typesetting and formatting characteristics, but it is compatible with computer-driven equipment, thus enabling us to break away from straight printouts or even computer output microforms as our only end product.

In 1974, when plans were laid to update the Commission's 1961 publication, Philip M. Hamer's A Guide to Archives and Manuscripts in the United States, it was intended to simply produce an update of this volume, following the same format, and including, even with the addition of all known repositories, perhaps 2,500 institutions. It soon became apparent that given the dramatic growth of institutions and the quantity of their holdings since 1961, the Commission's expanded mandate under the records program, the need for information about archives, manuscripts, and records beyond that included in the Hamer volume, and the availability of automated systems, that a different approach to the problem was needed.

The Commission's decision to make a commitment to a guide or data base as a continuing program mandated the need for an automated data system that would facilitate updating procedures, as well as providing the capability to produce a diverse group of publications. After evaluation of the potential of several systems, it was determined that SPINDEX II would best suit the project's needs. Subsequently, certain modifications, which will be discussed by Dennis Burton tomorrow morning, were made in the program, chiefly to increase its flexibility and ability to give us the variety of end products that we ultimately desire.

The Commission decided that the first step in developing a national data base of information would be the compilation of an institutional directory. Such a directory could be produced in a shorter time that it would take to compile detailed collection/record group information on a national scale, would fill an intellectual and administrative need, and would also enable the testing of the SPINDEX program package for the Commission's particular application.

In July of 1976 questionnaires seeking information at the repository level were sent throughout the country to more than 10,000 institutions that we had reason to believe had archival or manuscript holdings, which had been broadly defined to include not only traditional paper records, but also such variations as sound recordings, photographs, and computer data files. In response to this mailing and a subsequent follow up, we have received 2,700 completed questionnaires and 1,300 negative responses indicating that no archives or manuscript materials were held. Complete listings in the directory, including such information as hours and days of service, the availability of an institutional guide and/or other published finding aids, acquisition policies, and general restrictions on access, will appear for the 2,700 respondents, with abbreviated listings for 500 other institutions which have appeared in either the Hamer Guide or the National Union Catalog of Manuscript Collections, but which failed to return our questionnaire.

It may be interesting to note that in designing the questionnaire, as well as the basic tag structure for the project, individual elements appeared in the order in which they were intended to appear in the final printed product. In other words, the layout of the Directory volume itself was planned before the questionnaire and tag structure, with everything being designed in the same sequential order. This speeds up input as well as proofing time, and minimizes the necessity for cutting and pasting in the printed product to achieve the desired data order.

After verification and editing, information is entered via a CPT tape-typewriter device, which produces cassettes which are then converted into magnetic tape for processing. In addition to the basic repository information, the directory will include narrative introductory matter, a general index, and special lists of types of institutions, generated by selecting on a specially designed series of codes in the 49x tag.

Because of the capabilities of the system, future publications, in addition to updated editions of the Directory, may include a directory of acquisition policies, with its own index, or individual state or regional directories. The Directory is the first step toward the Commission's goal of a comprehensive data base of information on archives and manuscript holdings in the United States.

Our schedule calls for completion of NHPRC staff work on the Directory early this summer, and since we are holding to this work plan, it seems likely that the second phase of the project, development and implementation of collection/record group level data gathering, processing, and publication, will become a full-time staff concern at that time. Some preliminary developmental work was done by John Butler last summer, primarily in the area of analysis of existing collection/record group level guides and their patterns of description. As a result of this work, as well as our own consciousness of the needs of prospective guide users and what information could readily be supplied by institutions, a list of minimum and optional data elements was defined (Appendix B), as well as elements that we felt should not be encouraged. These requirements, it should be noted, are intended to apply to the NHPRC national data base; institutions using SPINDEX or any other system for internal control should not regard them as hard and fast guidelines for this purpose.

Minimum requirements for the national data base include such elements as repository name, collection or record group title, date span, quantity, and NUCMC number, while optional elements include such categories as a brief description of the collection/record group's contents, a description of the restrictions policy, if any, and indication of the existence of a finding aid for the material. Since no data has yet been entered into the system at this level, these requirements are still open for discussion, and we welcome comments and suggestions from the profession and potential users.

Because of the scope of the project and the complexity of the work involved, we feel that it is essential to proceed carefully, with thorough planning and testing of procedures by the guide staff at the NHPRC, and by a small number of projects at the state or regional level throughout the country. The one project that is currently underway in this area is a statewide survey in the state of Washington, and a careful evaluation of the results of this survey will enable us to provide more cogent advice to prospective projects in other areas. Many questions have to be answered during the development phase, such as the optimum level of entry--record group, subgroup, or series--for archival materials, the feasibility of constructing a national thesaurus and authority file, and the comparative cost and user effectiveness of data entries including minimum or optional elements.

We see the role of the central guide staff at the NHPRC as a coordinating and advisory one, and while we do foresee the probability of some type of centralized input of information from published sources, or other available materials, we also know that the national data base cannot be constructed without information gathering and actual data input into the system by state, regional, or multi-institutional projects in the field.

Like the Directory, collection/record group level publications can be issued in a variety of formats. There may be one general guide, issued on microfiche. There may also be special state, regional, or subject guides. Selection features and codes, which are being used in the Directory to produce lists of types of institutions, can be used at the collection/record group level to produce subject guides. It is also possible to issue a combined directory/guide publication. And like the Directory we are publishing later this year, the collection/record group data and publications can be updated and reissued at regular intervals.

In addition to the guide/data base project, the NHPRC is exploring other applications of SPINDEX to our administrative and intellectual functions. A beginning has been made in designing a data base that will act as a control mechanism for the records grant program, enabling us, for example, to determine the numbers of grant applications and funded proposals on a subject and geographical basis. It is conceivable that a similar scheme could be developed for the Commission's publications grant program, or for other Commission operations. The national data base and resulting publications, however, will probably continue to remain the Commission's chief SPINDEX application.

Lesser: According to my watch, we have all of five minutes for questions. I'll be happy to direct them to any of us who were speaking this morning, if anyone has a short question or two. Or comment? Nancy.

Sahli: I would just like to make one comment on something that George said. You mentioned that SPINDEX III would have an additional permuting feature, but I don't think that there are to be any additional permuting features; it's the selection capability that will enable you to do what you want there. Is that correct?

Hannestad: I believe so, because if what you were looking for were merely another field that could be used for permuting purposes, of course you could handle that in SPINDEX II by assigning a permuting feature to that field.

Mariz: I discovered that.

Hannestad: The real additional power of SPINDEX III is going to be in its selective and associative capabilities, and also in that it will move to a legitimate two level index, where SPINDEX II is very much held to a one level index.

Lesser: Any other comments or questions before we break for lunch?

Hickerson: Let's start back at 2:15.

SESSION II

Designing Tag Structures*

Brent G. Thompson: This afternoon, as you know, we are going to talk about tag structures, something that should be of interest to all of us. I hope that we'll have time for discussion afterwards. We're getting started a little late, but we should still have time to talk about some of our problems with tag structures and discuss the future.

Dennis A. Burton: "Tag Structure: NHPRC Guide"

The end products of the NHPRC guide project are publications--either paper or microform. Although desired output should always be considered first in a data base design, our project had to consider not only the sequence of data elements, but how they are to appear in a printed publication.

Thus, our first step was to design (in part) the books we anticipated publishing. Several months were given to this activity. When this process was complete we had the sequence of data elements.

We next evaluated the data elements to determine their tag assignment. The evaluation criteria were: 1) Would the data require frequent updating? 2) Would we want to select or index on that data? 3) Would a standard header be required? We did not want to use any more tags than were absolutely necessary.

By using the features of our print composition package we were able to incorporate several data elements in a single tag (i.e., tags 110, 120, 130, and 140), or to combine two tags into a single element in the final product (i.e., tags 310 and 320). The print package also allows the creation of standard headers for certain tags (i.e., tags 140, 200, 210, 220, 230, 24x, 300, 310-320, 33x, 380, and 39x). Some of our tags are reserved primarily for selection (i.e., tags 49x and 790).

*Papers by George E. Mariz, "SPINDEX II at Western Washington University" (Appendix D), Brent G. Thompson, "Guide Project Tag Structure" (Appendix E), and Joel A. Shirley, "Designing Tag Structures" (Appendix G) were contributed to this session but not presented verbally.

Although, in one sense, our file now has three levels (repositories within cities within states), we have treated all as level one to reserve more levels for detailed descriptions. The print package separates levels based on tag designations (tags 180 and 190).

Since each index term is assigned a tag (we do not permute), we often exceed the 50 tag limit of SPINDEX. Level 8 is used for this overflow.

The following tag structure is currently used for level one records:

100 Index reference (an alpha-numeric condensation of the control number)
110 Parent institution name(s)
120 Units within an institution
130 Street address
140 Mailing address
160 Telephone number
180 City
190 State
200 Hours of service
210 User fees (if any)
220 Access policy
230 Copying facilities available
24x Acquisitions policy
300 Quantity of materials
310 Earliest date of materials
320 Latest date of materials
33x Description of materials
380 Institutional guide titles
39x Bibliography

INDEX TERMS

42x Corporate or organization archives held by the repository
43x Geographic
44x Supplied subject
45x Acquisitions policy index terms (all)
46x Personal names
49x Type of repository code
790 Date of entry or update
800 Reserved for special index entries
90x Abstract

Dennis A. Burton: "National Data Base Design"

The NHPRC guide project staff was given the opportunity to present a proposal for a national tag structure, including all hierarchies. Although we have spent considerable time in working on such a structure, we feel that

it is too early to propose a data base design.

However, we would like to present some thoughts on this and related matters for discussion.

1) LEVELS

SPINDEX III will also be limited to 8 levels. However, unlike SPINDEX II, all of these levels can be indexed.

The guide project uses level 1 for repository information and level 8 for overflow index terms. We are considering suggesting assignment of the remaining levels as:

Level 2 - Record Group/Collection
Level 3 - Sub-group
Level 4 - Series
Level 5 - Sub-series
Level 6 - Box-Folder
Level 7 - Item

2) CONTROL FIELD

Although there are adequate levels available for an item-repository hierarchy, following the NHPRC guide format leaves only 9 control field positions for levels 2-7. By using an alpha-numeric design, most, but certainly not all, materials could be accommodated.

Another alternative would be to expand the control field.

3) INDEX SORT

The sort package in SPINDEX does not permit intermingling McC-Mac, O'B-Ob, or Saint-St. Is there sufficient need and/or desire for such a feature to justify the costs that would be involved in program development?

4) COMPATIBILITY

It is our opinion that design of a data base that will serve all needs of all users is probably impossible. However, we feel that a data base design could "grow" as specific needs arise. Compatibility could be maintained by clearing tag assignments through some central office.

Attached is an NHPRC guide work sheet for tag assignments through level 3 (Appendix B). Of the 24 assignments proposed for level 2, 11 were added to satisfy the needs of three projects. Additional assignments can still be made as the need arises.

David W. Hastings and John F. Burns: "Designing a Compatible Tag Structure for the Statewide Records Survey"

Burns: The Washington State Historical Records and Archives Project has as its first task the discovery, survey and description of significant public and private historical records and manuscripts in Washington State. Over 1,600 organizations and institutions will be inventoried including public agencies, private repositories and selected active private organizations. The product will consist of guides with descriptions ranging from repository to series level, and with subject, name, chronological and geographic indexes. It is estimated that approximately 10,000 separate collection or record group level entries and 18,000 series entries will be included. The need to arrange and index this amount of data necessitates computer control and the SPINDEX III system will be implemented to provide that control.

Working with the National Historical Publications and Records Commission and drawing upon the experience of other projects using SPINDEX, the project staff has selected descriptive elements which suit the needs of researchers, professional and amateur, public officials, and the requisites of the proposed national guide project. The survey format has been standardized for these elements in order to facilitate computer control, but as much latitude as possible is allowed in the substance of the entries in order to relate the distinctive characteristics of the records being described. The result is a balance which will enable us to conduct a comprehensive survey of the state's historical records in a short period of time, while still recording sufficient data to be of significant research value.

Data collection in both the public and private sectors begins at the institution level. For public records, data is then recorded at the record group level, occasionally sub-group, and series level, the latter being most important for state research needs. In the private sector description is at the collection level in all but a few cases which require data collection at the sub-group or series level for clarity. In order to integrate entries for both sectors into the guide the manuscript collection level is considered equivalent to the public record group level. The format for data entry is consistent for public and private records but the fields have been structured to allow for their diversity, and some aspects of data collection are treated differently for each sector. At the record group and collection levels sufficient compatability has been ensured to produce a guide which will direct a researcher to the appropriate repository and the specific materials necessary to his study, regardless of the origin of the records.

The project will be using the SPINDEX III system, scheduled for completion in the summer of 1978. The modification of SPINDEX II which will be incorporated into SPINDEX III will provide greater indexing capabilities, the capability of generating directories by accessing the control number, and greater composition flexibility. In addition, SPINDEX III will permit

the production of guides based on the contents of a tag field. With this feature it will be possible to produce directories to all records pertaining to a certain county, all records in archival custody, as well as directories to photograph collections, map collections, types of repositories, and many others.

The data collection forms have been designed to take advantage of these innovations, but since the modifications to SPINDEX are not yet final and since there is no documentation available, some changes in the format are expected prior to data input and no tag numbers have yet been assigned to the fields. Decisions have been made, however, on the substance of the data input and the basic structure of the fields, based on the capabilities of SPINDEX II and the information currently available on SPINDEX III.

Two distinct forms have been developed for collection of data, a repository form and a multi-purpose records description form. The repository form identifies each organization and institution surveyed and relates information about their collecting policies, if any, research facilities, and the general nature of their holdings. It also includes data entry fields for the control number, name of institution or organization, address, telephone number, days and hours of service, availability of copying services quantity and inclusive dates of holdings, restrictions on use, publications or places to which holdings are reported, availability of finding aids, type of organization or institution, county of location, date the information was collected, and an indexing field noting major subject areas, personal names and geographic locations. It would be redundant to discuss these fields in detail since, for the most part, the entries are generalized from the fields on the description form, which will shortly be discussed in greater detail, and because they are closely related to the NHPRC repository level form, which is widely familiar. Rather, the focus of this paper will be on description below the repository level (Appendix A).

Hastings: The description form is multi-purpose by design. It can be used for recording data at the collection, sub-group or series level in the private sector, and at the record group, sub-group or series level for public records. To accomplish this the descriptive elements have been categorized into discrete fields so as to facilitate data entry in a standardized format, while still allowing a degree of flexibility within certain fields.

The present format of the description form has been developed through the study of forms used in similar surveys, consultation with NHPRC, suggestions from members of the Washington State Historical Records Advisory Board, and from experiences during the first phase of the project. The format has already undergone several revisions, but the substance of the data collected has remained essentially the same. Most of the changes have been made in order to simplify and help speed the surveying process while

retaining accuracy in the essential descriptive elements.

Nevertheless, the form is subject to further modification. Assignment of tag numbers has been postponed and it is expected that some modifications of format will be necessary to comply with the requirements of SPINDEX III. Additionally, index terms will be supplied on a separate page appended to each form. For present purposes the description form includes entry fields for the control number, title, forms of material, quantity, dates, description, a few distinct characteristics, finding aids, restrictions on use, date the information was collected, other comments and indexing fields. The type of data entered in each field will vary according to the level at which the material is surveyed and whether the records are of public or private origin. Specific discussion of each field follows.

CONTROL NUMBER

The control number consists of five levels of entry comprising sixteen numeric or alpha-numeric characters, plus one entry for a record action code. The first eight digits will be assigned by NHPRC so that the data can be easily integrated into the national guide project. The final eight characters will be selected to best conform with the collection, record group, subgroup or series designation of individual organizations.

SPINDEX III will permit the production of directories by accessing any level of the control number. This will allow the production of guides to all of the repositories or collections in a given city, and will allow the generation of directories to the collections of particular repositories. Control numbers have yet been assigned to the data, pending possible revisions necessary for SPINDEX III. Space has, however, been included at the top of the repository and description forms for the entry of the raw data which will later be converted into control numbers.

TITLE

This field provides for the entry of titles for collections, record groups, sub-groups or series, depending on the data level. Allowance has been made for variant titles, especially useful in the public sector where the name of a particular series may vary with changes in administrations. In cases where the material is not titled, a concise descriptive phrase is substituted.

The field will be non-permuting since the titles do not always convey significant or even accurate information. Later, in the indexing process, subject terms and personal names will be selected from the title field to be entered into the indexing field.

CONTENTS

The contents field appears on the form as a group of descriptive terms with a box to be checked if a term applies. They are descriptive of the types of material and include such terms as "minutes," "reports," "correspondence," "plats," "ledger," and so forth. Allowance has been made for the addition of terms not listed in the field, and for the designation "see title" if the title is sufficiently descriptive of the contents. The field has been included to provide for a certain amount of standardization of description, to save time in the surveying process, and to decrease the length of the description field entry.

The contents field has been the subject of a certain amount of controversy among Board members and staff, as they represent various public and private institutions and several schools of thought in regard to records description. The current list reflects a good deal of compromise, and has been amended to include terms descriptive of the materials actually found in the first phase of the survey. The number of terms in the field could be much expanded, but since it is necessary for the surveyors to have the terms in mind when filling out the form, an attempt has been made to keep it as short as possible while still retaining a high degree of utility.

The greatest problem with this field has been the tendency of surveyors to simply note "see title" irrespective of the nature of the material. This tendency has been checked through more intensive recognition training.

QUANTITY

Quantity is recorded in terms of cubic feet, linear feet, volumes, and units of microform, as appropriate. When the quantity is expressed in terms of number of volumes it is often also expressed in cubic feet or linear feet so that the physical size of the collection can be appreciated regardless of the size of the volumes. The "units of microform" refers to the number of fiche, rolls, jackets, or aperature cards, which are specified in the medium field.

DATES

The two date fields record span dates and bulk dates, with provision for years and months. If the exact dates are not available it is possible to enter codes for "not specified," "no date," or "circa." If portions of the materials being described are missing, this is noted in the description field.

Initially, some of the surveyors had difficulty in recognizing what is meant by "bulk dates." The term frequently has little significance when dealing with an on-going public records series unless there are breaks in it. Rather, it was included to indicate the direction in which a collection, generally in the private sector, is skewed.

DESCRIPTION

Entries in the description field vary according to the level of data collection and whether the records are public or private. Entries for public records on the series level include a brief statement of the function, purpose and arrangement of the series, a description of the kind of information habitually entered in the record, and notes on variations over time. On the sub-group and record group levels the field records the functions of the office of record and summarizes the series that it holds.

In the private sector, the collections are described in terms of format and subject areas, with a listing of the significant series within the collection. If a collection is encountered that is too extensive to be accurately described at the collection level, series level forms are also prepared. Also, organizations such as private businesses often necessitate series level description as their records are more similar to those encountered in public agencies. In general though, less than six per cent of collections are inventoried to series level.

Information entered in another field such as contents is not repeated in the description field. This makes it possible to complete the descriptive entry with a few narrative sentences. To some extent the information required has been standardized, as outlined above, and a list of suitable general descriptive terms has been provided. However, a certain amount of latitude has been allowed to properly account for special characteristics, and a more comprehensive thesaurus must yet be developed.

Because the description requires a narrative entry, there is a notable amount of diversity, reflecting the perceptions of the individual surveyor. Some changes will have to be made in the editing process, but for the most part a small amount of subjectivity has proved valuable, and it is through the analysis of this field that quick evaluation of the quality of surveyor work is most easily obtained.

CHARACTERISTICS

The characteristics field notes whether the material consists wholly of original documents, reproductions, or a combination. The field is included primarily to simplify the description field entry, but with SPINDEX III it would be possible to generate a directory to all original materials (or all reproductions) by accessing this field.

STATUS

The status field consists of two check-boxes to indicate whether the material is "active or inactive." This pertains mainly to records not in archival custody and it is especially useful on the series level in the public sector and in active private organizations, where the current activity of the record may not be readily evident. Apart from being descriptive, the

field could be used to generate directories useful to the State Archives for the regulation of public records.

ARCHIVAL CUSTODY

The archival custody field is included so that inactive repository directories can be generated and incorporated in the listings of the national guide project. It will also serve to identify collections which should be, but are not, in archival custody for future preservation efforts.

RESTRICTIONS

This field denotes restrictions on the use of the given record described. Often selected records have restricted use even though the repository or organization has no general restrictions policy. With public records, some of these restrictions are mandated by law.

FINDING AIDS

Since systems of finding aids vary greatly from one institution to the next, no attempt has been made to incorporate in-house location codes and the provision of finding aids has been left to the repository. The kinds of finding aids available, if any, are noted for the information of researchers. The field could also be accessed later to determine the state of intellectual control for future consultation projects.

DATE INFORMATION COLLECTED

This field will serve to inform researchers of the currency of the information to the guide. It will also be important as the guide is updated, since updating will probably be somewhat sporadic. In years to come the field could be used to generate updating priority lists.

TYPE OF INSTITUTION

This serves to identify the type of organization, political unit or repository and, apart from being descriptive, it could also be used to produce directories to particular types of repositories, or for statistical analyses of the volume and number of collections held by each category.

COUNTY OF LOCATION AND COUNTY OF ORIGIN

Here again, the primary purpose of these fields is descriptive, however they will also serve to generate guides to those records pertaining to a particular county regardless of the location of the records. This kind of directory will be especially important to the study of local history and genealogy.

MEDIUM

The medium field is also primarily descriptive, identifying the physical nature of the material. However, it will also serve to generate guides to photograph, map, microform, ADP, and audio-tape collections. If the medium of the material is microform, the type of microform will be noted, clarifying the figures in the quantity field.

COMMENTS

Entries in the comments field vary according to the data collection level and whether the records are public or private.

For public records, the field may include comments on the physical condition of the records and storage conditions, if unusual, as well as anything else out of the ordinary. In the private sector the entries may show biographical information (if available) and other information necessary to establish the significance and origin of the material.

In spite of instructions to the contrary, it is in the comments field that surveyors tend to be most subjective, expressing their personal feelings about an organization and its personnel, the project in general, and the more disagreeable aspects of their work. This sort of information can be enlightening, but it is in this field that much editing will be required. Data from this field is expected to be entered only on rare occasions.

INDEXING

The indexing field will be prepared as a part of the editing process. Indexing terms will be selected from the title and various other fields. The first step will be the preparation of a thesaurus. Existing thesauri will be used, but the final version will strongly depend upon a review of representative blocks of data from Phase I of the survey. Index terms may include public record series titles, personal and organizational names, geographical locations, and specific subjects. The object will be, of course, to direct researchers to their precise areas of study and to the records location and origin.

As mentioned above, the project is in the first stage and no data has yet been tested. Consequently this is all somewhat tentative. Some problems were experienced with the initial forms and necessary revisions have been made. No doubt further changes will occur.

The forms though have presented far fewer difficulties than those arising from having up to fifty different people gathering the information. Variations in perception and style of expression have made it necessary to closely supervise the surveyors beyond their initial training in order to

achieve the proper combination of accuracy, standardization and expressiveness. For the most part, the problem has been confined, but editing will be required to modify some of the more subjective and bizarre entries which still occasionally appear.

As yet, obviously, no comment on problems of data entry can be made, as implementation of the SPINDEX package lies ahead. However it is fundamental to the timely cost-effective success of any computer-based endeavor that considerable thought be given to the process of assembling data prior to entry. This has been attempted here, and it is only in forthcoming years that the success of the endeavor can truly be measured.

Thompson: Those of you who have worked on tag structures know what hard work it is. It's a very difficult process. The first question I would like to pose to the group is: Is it possible to have a national tag structure? I invite comments at this time on that question. Max.

Max J. Evans: I think that there are a couple of questions that are involved. First of all we need to identify some base, some common denominator. If we can do that, I think, yes, it is. Beyond that, how each repository or regional project wants to handle their tags, I don't think it matters. I don't think they need to check with some other place to see what tag they're using for county designations. Each can use their own. The other question is: If we have a national tag structure, how much can it be modified? The sequence of tags is important in SPINDEX II. How important will it be in SPINDEX III? I don't know the answer to that yet. If you come up with a tag sequence that I don't like for my own repository, I'm not sure I'll use it. But if it doesn't matter when we get into SPINDEX III, then that's not really a critical issue.

Stephen Hannestad: For a quick answer, SPINDEX III is going to maintain the same sequential tag requirements that are in SPINDEX II.

Thompson: Yes.

Frank G. Burke: It seems to me that when we talk about establishing a national tag structure, and considering the limitations that Dennis has put down as the way the thinking goes at the Commission, I think you have to be careful

about thinking specifically, or thinking generically. If you're thinking generically, as was brought up earlier today, then it doesn't become such a problem. We merely say, as on Dennis's list, that tag 300 is a quantity tag, and we don't care what that quantity represents, or that tag such-and-such is a title tag, and we don't care what title you put in there. Depending on what level you're in, all that we're talking about is information which can be compatible. If everyone is putting titles of the same things in, or quantities of the same things in, the program is then compatible. If everyone was putting in at level 1, on archival collections, the title of the record group, or the collection, if it was a manuscript collection, it seems to me that those who did that would be compatible with each other if they ever tried to merge their files. But, if Parmenter was putting in the title of an automobile registration in Montreal, then it would make no sense to merge that file with anyone else's. The only time that there is even going to be talk about compatibility is to the lower tag levels, and especially if you get down below level 1, or even level 2, then you start getting into variations such as shelf location or building location, material on campus, if you have your collection at more than one place, or donor, or whether they are in Hollinger boxes or in file cabinets, and all these modifications which are local options only, and of local use only. There, I would not like to get into an analysis of every possibility, every combination and permutation, and assign an irrevocable tag to it. I think this is what Dennis is talking about. If we can agree that at all levels 110 is a title, and so-and-so is a quantity, and such-and-such is a date, then that's all that we're really after.

Thompson: Stephen, you had a comment?

Hannestad: On the general subject of compatibility, compatibility is fine, provided it is geared toward some kind of exchange of data. But if you have an application where you know that there's no intent whatsoever to exchange data, then there should be no reason, that because standards exist, that you should not be able to continue to exercise whatever flexibilities the system has for your particular application. Standardization, in the absence of a reason to exchange data, becomes meaningless. Five, six, seven years ago, we attempted to dream up a standardized tag structure system but provided the users no rationale for using it. We had no existing data base for people to contribute data to. As a result, a standard is printed in the documentation, but that standard now serves only as an example of how tag structures can be used because there was no intent to exchange. There's no sense in having standardization unless you have an intent to exchange, and I think that each project that realistically envisions exchange is going to have

to develop its own project-oriented standard. For example, the steel industry decides that they want to start exchanging data about a particular subject in SPINDEX format, I would imagine that they, at that time, are going to design their standard tag structure. But until there is a need for standards, they should not be arbitrarily imposed.

Evans: But isn't there, at some level, some fundamentally identifiable fields that we could assign tags to? Beyond that, of course, each individual repository can do their own thing and assign their own tags to other fields that they need, but is there a central core that can be isolated?

Hannestad: There's a minimum central core. You do not have a choice as to where you put your control field. It has to be in 000.

Evans: That's pretty minimal. (laughter)

Hannestad: You can not, without producing a very weird printout, redefine the 150 tag because that triggers a very special "see" routine or "see reference" routine in the index program. It does not pay "to mess with mother nature" and mess with the 400 tag, because the 400 tag has a parameter card insertion option. So, right now there are certain system-oriented, reserved fields, and, indeed, there could be more fields which could take on a reserved meaning. But I would like to see some system-justified reason before we start to reserve any fields.

Thompson: Let me bring up a personal example at this point. In our repository, we use the 120 tag for the creator of the collection, in library terminology, the main entry of the collection. As far as I know, other repositories don't appear to do that. So that is a fundamental change which we did in our tag structure which other people don't do. At this point in time we probably wouldn't be willing to change from that, and other people may now be in the same position or may later get into a position where they have a basic difference with the NHPRC tag structure. Frank.

Burke: I guess there are really two questions here. One is that when the initial tag structure was designed, and 100, 110 and 120 were reserved, that was an old format that in SPINDEX I was card A, card B and card C in effect, or there was a code in the card that was a number 1, number 2 or number 3 that implied that those fields did certain things when you went to index them. Tag 100 didn't do anything; it didn't get into the index. Tag 110 permuted, and tag 120 alphabetized the entire field and listed it

that way. So there was a purpose to that. That purpose was done away with once you got into the parameter card control for making any tag do anything you want it to by throwing in a parameter card and making tag 380 act as though it were 110. So it took away the purpose of structuring it that way but when we got to the guide itself, we are getting to the point, that Steve has mentioned, where there is a rationale for everyone trying to get together and decide on what the system should be, so that if you wish to feed us information, we can get it into the national guide without having to restructure it. But there is the possibility, is there not Steve, that, if the material came in from Salt Lake, and the 120 tag should really be a 100 tag in our system, and if we were processing only that data, we could throw in a parameter card in order to treat that 120 tag as if it were a 100?

Hannestad: It would not be quite that simple. What would be necessary would be to take their data and move the fields, doing a convert on the tag number. It can be done automatically. Certainly it would be simpler than replacing their data; you know, rekeyboarding it. But it is not something that can be done strictly by parameter card. The parameter card will change the indexing function of the field, but, whereas 120 for you means units within an institution, and 120 to them means something else entirely, we would have to change their tag number and restructure their record. However, again, this brings up a very fundamental point; as long as each user documents what is, in fact, in his file, the computer can make the adjustment without manual reinput of the data. What will become chaotic is when you leave no trails behind, and you don't tell anyone who gets your tape what field contains what data. It's that lack of documentation of the data file which will cause the ultimate chaos; not the fact that three different institutions use three different numbers to describe the same concept.

Burke: But, that's only a retrospective concept. I mean, if we're talking about cooperative programs, then we're not encouraging people to use any field that they want because the computer can take care of it later.

Hannestad: No. Certainly all institutions participating in the national guide project should adhere to a certain minimum level of consistency between each other. But, if somebody wants to put a firehouse response system into SPINDEX to tell what fire engine responds to what fire, there is absolutely no need. As a matter of fact we have done that with SPINDEX.

Burke: A firehouse response system?

Hannestad: Yes, a firehouse response system to tell you what apparatus to send and what special hazards may be encountered at the fire. Compatibility should exist only where compatibility is useful. If NHPRC

guide project members all use the same compatible tag structure, fine, but that does not imply that that tag structure is a gospel that would have to be used by Coca-Cola, or that it would have to be used by STELCO [Steel Company of Canada, Limited]. And, in fact, I can imagine that within a particular industry such things as technical literature files might form a basis of exchange between STELCO and International Nickel [INCO, Limited]. They do talk to each other?

Thompson: Tom.

H. Thomas Hickerson: You know, I think that all that is obviously true, and among some applications in your repository, you may not have any expectation to mix data. In doing our university publications, we didn't think we would ever mix it with university archives collection indexes. Yet, because we didn't really know, we tried to make it as compatible as the two structures could be, just for potential. The thing is, there is no reason for us to do it differently if we can do it in a relatively compatible way. So people can always have freedom to do anything they want to; I mean, that's not the question at all. But, I think that we should have a common, suggested pattern. It doesn't place any limitations on anyone for that to exist. I think we have, perhaps, some problems on levels, as well. Would they be as easy to change? I guess you could index from several levels if people had coded their data at different levels than the NHPRC is using.

Burke: You can go up or down a level.

Hickerson: That wouldn't be any problem. I wonder how useful reserving a level for sub-series is. What was the reason for doing that?

Burton: Two or three people thought that was important. Let me say something, a little background; when we started designing our data base, we did it only to serve our own purposes. We weren't thinking about anyone else. At that time we assumed that we would probably never be interested in anything below the record group/collection level. Later we found we would want to go down to sub-group because you couldn't get meaningful description at the record group level. That was one assumption in the beginning, and then people started coming to us and saying, "Well gee, we would like to do what you're doing, however we've got to have a series description in there. How can that be done and still be compatible with what you are doing?" I think it goes without saying that anyone who has the SPINDEX package can do anything they want to do with it. They own it. It's not something they have on loan. On the other hand, is it desirable and possible to come up with minimums, a basic structure for those who want to be compatible with other people, so that they can be, and at the same time give them the

freedom to do other things that they want to do differently?

Burke: I don't think there's any doubt about it.

Thompson: Steve.

Hannestad: There's another advantage that comes from having certain standard base designs; the success or failure of any SPINDEX-oriented project is predicated on what amcunt of thought goes into a data base design. To the extent that a project can pick up and use an existing standard data base design, they then save themselves the entire trauma of coming up with a new data base design. It should be stressed that if you make a mistake in your data base design, you are stuck with it throughout. If you put three different, completely different, concepts of data in under a single tag number, the computer will never be able to sort those three out as being different concepts, because it's going to see a tag 573, and it's only going to be able to catch that as a category 57x piece of data. If you later on decide you have to add a new type of data to the field or to the record, that can be done without too much trauma, but if you have already mixed apples, oranges and pineapples together in a data field called fruit, then fruit is all you'll ever have. There's such a desire when you get SPINDEX, and you've heard all the good things it will do for you, to run to the keyboard and start putting data into your system. Yet, in most cases, you have to sit back and wait several months thinking about what you want the final product to be, and then working backward from the final product to identify those discrete data elements that are necessary. Once the discrete data elements have been defined, numbers can be assigned in a matter of minutes.

Burke: I think maybe we can wrap this up by answering one of the questions that Dennis posed. What we came here with was this rough sketch and, I guess, a desire to get an indication whether we should proceed on this premise, that we attempt, down to a reasonable level, to establish what we feel are the appropriate fields, associated with appropriate tags, and then promulgate that among all the users of SPINDEX, saying, "Here we are. Beyond that you're on your own. If you stick with that, we will all be compatible to that level." The suggestions which we have put down are those we have thought about so far, and thought long and hard about, and that's where we are. If everyone agrees, then we'll go off and do a little more work on it.

Hickerson: I think that is a good idea; at the same time, obviously, as Dennis just said, you designed the tag structure for your project.

Burke: True.

Hickerson: At this point, there are probably more common elements among some of the other users' tag structures than there are with the NHPRC tag structure. I guess my two questions are: What sort of mechanisms for review will there be, and what is the possibility that we not sign on to your tag structure, but instead, you modify yours to match points of agreement among other users?

Burke: I guess we would have to see what the situation is before anybody can make any decisions as to what's wrong with it, you know, what needs to be done.

Thompson: I think I would agree with that. It is difficult at this point, not knowing what SPINDEX III is going to do and not knowing what other users have, to decide whether this particular tag structure can meet the needs of repository guides for ourselves or Cornell or South Carolina and also meet the needs of the national guide project. Personally, I don't think we've been able to answer that question so far.

Burke: I don't think III is going to make any difference really, because, first of all, nobody is really in the format of III yet, therefore, nobody is stuck with it, and it's an advance on what II is. Anybody who's got II will not be outdated by the appearance of III. It is not going to wipe it out. You're not going to have to rekeyboard it.

Hannestad: There again, I must stress that III is maintaining the concept of sequential tags. So, there's no reason to hold up any standardization of tags waiting for III, because III is going to have the same tag features that II has.

Thompson: So, in other words, no matter what my repository's or any other repository's tag structure currently is, SPINDEX III will not alter that.

Hannestad: That is correct. The only difference will be that in the index you will be able to call one tag out of numeric sequence and throw it to the front of the line.

Evans: But, because the printing sequence, with that exception, is based on the tag number, that makes it more critical that we all get our heads

together and decide what order might be best for each of us and see if we can come up with some sort of agreement and leave enough holes in the numbers to fit in things beyond the basic, fundamental core of fields.

Burke: I guess what we need is for everyone, who has a tag structure and is using it now, to see everyone else's in order to see what all those tag structures are and to make comparisons.

Evans: What we need are the fields and the order we want them in, and we don't care what the numbers are; I don't anyway. They don't mean anything to me except for the fact that they're sequential.

Charles H. Lesser: We have in our packets ours and Brent's. Is yours in the packet somewhere Tom?

Hickerson: No it's not. It's available in the report on our application, but it's not in the packet.

Thompson: I'm not proposing that whatever the LDS Church Historical Department has done should alter what the guide project is doing. I'm asking whether other repositories (and I hope there will be others), when they begin developing tag structures, can be compatible?

Richard H. Lytle: Just a point of information for some of you that might not be acquainted with it, there is a philosophy which represents the opposite of what we're discussing here, the freedom of SPINDEX to design your own tags. I'm not advocating it, but it might be of interest to you if you're concerned with these problems. The Museum Computer Network has adopted the complete, total reverse, wherein a single organization rigidly describes what a tag will be, and tags are conceptually, very narrowly defined. Everyone in the Museum Network must use a tag exactly this way. You might write them and ask them for their latest, little manual which gives you all the tags and all the concepts. I'm not arguing for that; I'm just mentioning it as an interesting contrast. I don't think the tag situation is as simple as it's being described here, somehow. By that I mean, it would be fairly simple to make basic standards, and then everything would just follow along. I think it is something which deserves a good deal of study and thought. At the very least, to define the minimum which ought to be collected everywhere deserves a lot of thought.

Hickerson: There's a lot more to the apples and fruit question and the specific definitions of descriptors we use, even a descriptor like a personal name tag. We intermixed personal names of correspondents and personal names of subjects of correspondence in the same field, and certainly at this point, there is no way we can economically go back and change that data. There are common problems, like how people apply the definition to the data to be gathered for that tag, that enlarge the problems of tag enumeration and defining the data to be entered in the tags.

Lytle: That's why I suggest taking a look at that Museum Network thing, because you can see requirements of tags and tag definitions. They've got metaphysicians working. (laughter) This is conceptual, Frank. They've got tag 112 reserved for describing a statue as viewed from a perspective of four feet, slightly to the left.

Thompson: I'm afraid we had probably better break it off at this point. I hate to do that; I know there are unanswered questions. But, if we're to get through the program, it will be necessary.

Burke: Can I make a closing statement?

Thompson: Okay.

Burke: I hate for things to be left up in the air on this. Why don't we just proceed by an analysis? If everyone who has a tag structure can just send us a copy of what their tag structure is, we have two of them here, we can then proceed to an analysis of those and the proposed structure which Dennis drew up, that you all have, and come out with something we can then circulate among the users and get some further discussion of it, either next OAH [Organization of American Historians] meeting or something like that, or by mail. At least we can move forward on something and make a suggestion, but we are going to have to come down to the point where, if it isn't agreeable to everyone, for our own operation, we're going to just have to go this way.

Thompson: Yeah, it's obvious that you will have to decide what is best for NHPRC.

SESSION III

Data Entry and Computer-Generated Results

George E. Mariz: This session focuses on data entry and the generated results, and in it we have three separate papers. The first is by Joel Shirley of the South Carolina Department of History and Archives.

Joel A. Shirley: "Data Entry and Output at the South Carolina Archives"

The South Carolina Archives has been working with SPINDEX II since July 1971. We have actually been entering data since July 1973. Indexing information for over 100,000 archival documents has been encoded; finished indexes to 80,000 of these documents will soon be available on computer output microfilm (COM) in our Search Room. COM indexes to smaller numbers of documents have been available to researchers since 1975.

To get this information into the computer, we are currently using a Computer Machinery Corporation CMC-3 keyprocessing system. The CMC-3 is a batch-oriented, software-controlled key to disk system. We have two full-time data entry operators who use two of the CMC-3 CRT (Cathode Ray Tube) keystations for data entry and verifying and a third CRT station is used for reading data while on disk. The disk has a capacity of 18,000 80-character records. The CMC-3 is leased on a 3-year contract basis for $580 a month (maintenance included).

The CMC allows us to keep the input data in a manageable form (batches of about 100 input sheets). Batches that have been encoded and verified are transferred to tape by the CMC unit and are sent to the South Carolina Division of General Services for listing. After the paper list has been edited by an archivist, only the batches needing corrections are reloaded into the CMC unit and only the specific records that require correction are recalled to the CRT station.

Perhaps this is the best point to speak about editing. Even though input sheets are checked before entry, thesauri are periodically updated, and entered data is verified, every batch of records is edited by hand at least once after it is encoded. Most errors are corrected at this time. We also carefully check the ARSP4 "Diagnostic and Error Listing" for errors that we have not otherwise caught. The fact that only eight significant errors have been detected on our finished products is indicative of how fussy we are about the accuracy of our end product.

We have thus far used two basic methods of data entry: input sheets and direct data entry from existing indexes. Input sheets are made out for series for which no adequate indexes currently exist, for series for which existing indexes are in a format that is not easily converted to our SPINDEX

format by encoders, for series with indexes that are not compatible with SPINDEX, and for all other cases in which an archivist is needed to make indexing decisions. Data entry by this method is, of course, more expensive than direct entry from existing indexes, in that it requires paying the salaries of archivists (see example 1).*

From our experience, an archivist can index from scratch between 20 and 130 documents a day (depending on the depth of indexing, length of the document, etc.). If at all possible we try to index a series at the same time as other processing activities are being performed (such as arranging and foldering unarranged series after the restoration of the paper documents involved). The cost of filling out an input sheet and editing can be as much as $5.00 per document, which does not include computer processing costs. We are therefore very selective of the record series that receive this attention.

The second method we use for data entry is entry directly from existing indexes (see example 2).* If an existing index fulfills certain criteria (i.e., has all the necessary data elements for our computer output index, needs very little interpolation of additional words by the encoder operator, and has easily determined document numbers), then the data entry operator enters the information directly from the index. This bypasses the archivist except for editing. One series of colonial land grants containing 26,000 documents was recently completed primarily by this method at a cost of only 45 cents a document or about $12,000.

One other method of data entry we are beginning to investigate involves using archivists as data entry personnel. This is an attempt to merge the best features of the two previous systems: maximum performance at a minimum cost. In an ideal situation this would best be done using a CRT screen that would show all or most of the indexing input for a single document at one time and be more or less a "fill in the blanks" operation. We will test this with our current equipment in the very near future.

Our present rate of entry, using the two implemented techniques, is about 3,500 documents a month. To achieve that rate we have three archivists and two data entry operators working full-time on the computerized document index and often have an additional part-time intern. Quite a number of other staff members in the department prepare input sheets as a part of their regular work. Arguing that these other staff members (and indeed the archivists working full-time on the project) would be involved in some other more traditional effort to gain control of our holdings if we had not adopted our SPINDEX system, we have not included a portion of their salaries in the calculations we have made of our per document indexing cost. Over the last year our cost per document including the salaries of the staff working exclusively on this project, computer time, encoding, and output averaged $1.13.

* Appendix F.

Getting the data into a SPINDEX format, encoding it, editing printouts of the data, and building master files are the major phases involved in building our data bank. To utilize that data bank we have chosen for the present to produce document indexes on 16 mm computer output microfilm. It is conceivable that sometime in the future the size of our data file and the volume of use might justify the expense of on-line access.

COM has many advantages over output printed on paper. For one copy, COM costs one-fifth of the amount we pay for paper. In most archives, libraries, offices, etc. space is at a premium; COM takes up only one percent of the amount of space occupied by the same volume of paper. The same proportion holds for weight. Film duplicates are also cheaper than paper, costing only 1/10th of a cent per page. COM is also a more permanent means of retaining the indexes than acid and flimsy computer printout paper would be.

We decided on 16mm roll COM instead of fiche primarily because we were already set up for roll film and most of the researchers and all of the staff are familiar with roll film. Roll film is also easier to keep track of than fiche. The roll film we produce is designed for 1,750 frames per roll with a title page at the beginning of each roll.

The University of South Carolina Computer Center has a Bell & Howell model 3700 COM unit which supplies our COM needs off-line. Microfilm originals are positive, silver halide; copies are negative, vesicular film. Originals, which we store separately as security copies, are supplied at a cost of $5 per 1000 frames. Copies for Search Room use cost $2 per 1000 frames.

The South Carolina Division of General Services performs all of our COM formating by the addition of the CM-360 software package into the SPINDEX job stream print programs (ARSP5 and ARSP8). We still maintain an option of either paper copy printout or COM for most of our input. For series comprised of loose papers we produce a paper register to use as a shelf list to enable us to periodically check to see that all the documents are present and in their correct location.

Even though we believe we can further improve our efficiency, on the whole we are satisfied with our data entry and output methods. The cost of the type of indexes we are producing would be prohibitive without SPINDEX. Of course, there is the question as to whether we would be attempting this high a level of access at all without the use of the computer, but as the size of our data bank grows and as it becomes possible to locate many of our most heavily used documents through one combined alphabetical index, we believe that an argument for SPINDEX could be made in terms of greatly increased staff efficiency in answering queries in addition to the more obvious argument in terms of a revolutionary improvement in intellectual access to our holdings.

Brent G. Thompson: "SPINDEX II as a Card Catalog Replacement in the LDS Church Archives"

Ron Watt has already explained how the LDS Church Historical Department became involved with SPINDEX II and the current applications we have made of it. Rather than talk about all of our SPINDEX II projects, I have decided to limit my discussion to our use of SPINDEX II as a card catalog replacement and the overall cost of SPINDEX II to the Historical Department.

The LDS Church Historical Department was neither the first repository to consider SPINDEX II as a means of replacing the card catalog nor the first to consider it as means of producing a guide to one's holdings. Cornell University and the Hoover Institute of War, Revolution, and Peace are at least two who conceived the idea much earlier.[1] We have, however, in a little over a year been able to produce computer-generated descriptions of over 5,000 manuscript collections, church record groups, and photograph collections. Quality SPINDEX-generated output can be produced in four basic steps: physical processing, data entry, data processing, and editing of output.

The first step, physical processing is the most important. Neither SPINDEX II nor any other present or future system of archival automation can help the repository which neglects this step. There is no substitute for well-done arrangement and description. Our description process includes filling out a worksheet each time a collection is processed. The worksheet requires the cataloger to supply main entry, added entries, supplied subjects, a collection number, a content note, and other useful information. We have found an authority list crucial to maintaining control over supplied subjects. It is just as important with SPINDEX II as with a manual card catalog. We also find it beneficial to provide full name and birth and death dates for personal entries whenever possible. Before we began to use SPINDEX II, the work sheet was used to produce cataloging cards. Then in June of 1977 the card catalog was closed. This required the cataloger to add a control number to each new worksheet together with the necessary three-digit SPINDEX tags. Obviously, our cataloging procedures have been changed very little. After a worksheet is filled out, it is checked by the processing supervisor for cataloging and coding errors.

The coded worksheets proceed to step two, data entry. Our data is entered on a dual station IBM 3742 which uses magnetic diskettes for storage. Its advantages are quiet operation, easy correction of errors (return to error and re-key), short training time, and low rental cost. The disadvantages are its upper-case-only input, lack of editing features, and small display screen. Prior to our involvement with SPINDEX II, we had leased the entry device for automating various name indexes. Our current cost for the IBM 3742, which SPINDEX occupies half-time, is $214.00 per month.

[1]H. Thomas Hickerson, Joan Winters, and Venetia Beale, SPINDEX II at Cornell University and a Review of Archival Automation in the United States (Ithaca, New York: Department of Manuscripts and University Archives, Cornell University Libraries, 1976).

Ron has already explained how we previously used a magnetic-tape typewriter to type a master card for each collection with the typewriter automatically producing the additional five or six cards. When we quit using that system in May 1977, we had an eight month backlog of untyped cards, our average for several years. SPINDEX cuts entry time to such an extent that we were able to input the eight-month backlog in six weeks with no increase in clerical staff time. Our clerical personnel can enter twenty-five worksheets an hour, the equivalent of 150 cataloging cards, and further clerical time is saved since the tedious and mistake-ridden process of filing cards is eliminated. We have saved four hours of clerical staff time each day by switching to SPINDEX II-generated catalog descriptions. We use the saved time to input data for other SPINDEX II projects.

After the desired data has been input, the parameter cards are selected. One of the advantages of SPINDEX II is the ability the archivist has to control computer output by modifying the parameter cards. He does not need to rely on an expensive programmer's time who will not understand what the archivist wants. We have established a set-deck of parameter cards which are sent to the computer center with every guide project job.

The diskette or diskettes are sent with the parameter cards and job instructions to the computer center for step three, data processing. These diskettes are converted to tape and the tape is used in producing the desired output. We produce four different printouts. The transactions diagnostic (ARSP4LST) is a listing of new records used to update the master file and records which, because of errors, could not update the master file. The second printout is the current master file (ARSP40). It provides a complete list in control number order of the updated master file. We call the third printout the guide (ARSP51) The guide is a formatted list in call number order of our cataloged collections and it includes main entry, title, collection number, and content note for each collection. It is produced on white paper, burst, trimmed and placed in binders. The index (ARSP81) is a list of main entries, and subjects arranged alphabetically with the appropriate titles and collection numbers appearing under each keyword. The collection number refers the researcher back to the guide where, by reading the content note, he can find out more information about the collection. The index is prepared in the same manner as the guide.

A price breakdown for our last printout might be helpful. The transaction diagnostic, which contains 1,241 records and occupies 191 pages, costs $40.48. This expenditure includes updating the master file. The master file cost $51.57 for 4,677 records and 1,046 pages. The guide also had 4,677 records which appeared in 613 pages and cost $31.93. The index which is the most expensive part cost $75.87 and contained 11,989 individual entries on 391 pages. The total cost for all four parts was $199.85. We produce an additional index with the same master file but with different parameter cards called special lists. Some of the items included are lists of journals and diaries, non-English language materials, daguerreotypes, restricted materials and donors.

The last one produced contained 3,296 individual entries, 88 pages, and cost $16.50. In the manuscript guide we have the equivalent of 24,139 cards under our old card catalog system. We are considering using microfiche for the guide and index which would save money when multiple copies are needed.

The editing of the printouts is step four. The diagnostic is reviewed both for those records which, because of errors, are not added to the master file and for new additions which are proofread. The needed corrections and additions are then noted on the master file printout. The index is checked for input and cataloging errors with the corrections being marked on the master file. The index is much easier to edit than a card catalog for two reasons. First, the archivist can scan through a printout more quickly than he can thumb through cards in a catalog drawer. Secondly, common terms and names are listed under one heading making misspelled names, subject headings, or incorrect dates conspicuous. These advantages coupled with SPINDEX's easy system of making corrections make our automated finding aids much more accurate than our old card catalog system. After the guide is edited, the master file printout with all its corrections, additions, and deletions returns to step three where the clerical staff input the changes. We input new records and correction records together and find no problem in doing so. We plan on updating the guide every three months.

Many archivists seem to fear the total cost of the SPINDEX II system. We feel that the advantages gained by SPINDEX II justify the costs. Because computer firms have automated accounting records, I have been able to find out exactly how much money we have spent on our system. It is impossible to separate the guide project from our other SPINDEX projects because many costs apply to all projects. However, the replacement of the card catalog is our most important project and over 90% of what we term "production costs" have been expended for it.

We purchased SPINDEX II in late 1976 for $500.00 and then spent $1,424.32 to set up the programs. Included in that cost was writing a program so that data input on the IBM 3742 could be accepted by the card loader (ARSP1A) program. During the developmental phase between December 1976 and May 31, 1977, we learned how to operate the SPINDEX system and were able to develop a tag structure which allowed us to replace the card catalog. We spent $1,533.96 during this period. There is much frustration involved in experimenting with the programs. For many discouraging weeks every test run resulted in an abend. By June 1, 1977 we had either solved or knew what needed to be done to solve our problems.

During June and July we completed major modifications to the programs. The ARSP6 programs were modified in such a way that index entries would appear in an order closer to cards in our card catalog. These modifications were quite expensive with the cost for programming and computer time being $2,226.30. We were then able to go into production and spent $5,922.04 between June 1 and

December 31, 1977. During this time we coded worksheets which were already in the card catalog and were able to enter most of the data for our guide project. We also produced a microfiche guide to the department's oral history program. This year we have spent an estimated $1,400.00 on production. All totaled we have spent $12,106.62 on SPINDEX II. I have not included the cost of our entry device since we'd leased it previously for other purposes. However, the prorated costs for SPINDEX's use of the IBM 3742 would be $1,605.00. I have spent approximately 80% of my time working on SPINDEX II during the fifteen month period.

The costs of SPINDEX II must be placed in perspective to the overall cost of archival work. The historical department has twelve professionals, four paraprofessionals, and three clerical employees whose job is to provide for the intellectual control of our holdings. Costs for staff, floorspace, supplies and equipment make the cost of SPINDEX II seem small by comparison.

Furthermore, other users will not necessarily find the same costs since costs depend upon amount of data to be entered, hardware costs, CPU rates, etc. They may be able to spend less with the same results by using the standard programs without modifications or by having less startup and developmental costs.

The Historical Department has found SPINDEX II able to produce a quality card catalog replacement at a reasonable cost.

Richard S. Parmenter: "Data Entry and Computer Output Methods"

Data entry is a particularly important phase of all SPINDEX applications. Each new application brings with it a tremendous backlog of information which must be converted to machine language. There is no question that data entry is the single largest and most continuing "bottleneck" in the SPINDEX operations.

Since SPINDEX applications are primarily in clear text as opposed to code and use qualitative rather than quantitative descriptors opportunities for improved cost effectiveness abound. Inco now has fourteen years continuous experience entering data for information retrieval index production. Through systems analysis techniques, improvements have been readily identified.

INPUT METHODS AND EQUIPMENT

IBM Key Punch
IBM Diskette Inscribers
IBM Mag Card Typewriter
IBM Typewriter Terminal, 2741, (ATS Shared Logic)
Datapoint Intelligent Terminal

We employ at least five different variations to enter data into the SPINDEX system. Punched cards are usually the initial system employed. Even where improved methods have been adopted, card entry job streams are maintained. Peak loads are frequently done on card punching equipment in-house or at a service bureau.

Several locations have replaced standard IBM keypunches with IBM diskette inscribers. Floppy disks are usually converted to standard nine-track magnetic tape for entry into SPINDEX.

The other three methods all involve terminals which either transmit entry data to the main computer, or prepare magnetic media which is converted in a separate operation for use in SPINDEX.

For a long time in New York, we used the IBM Communicating Magnetic Card Selectric Typewriter. Magnetic cards were converted to nine-track magnetic tape which went into SPINDEX. In one plant, we use typewriter terminals with no recording media through telephone lines to the computer.

In New York, we are testing and developing our latest method of data entry--the multipurpose intelligent diskette terminal. We are currently using the Datapoint 1100 two-diskette terminal with a printer and nine-track tape drive. We have just completed two years of operation and have plans to upgrade this equipment.

OUTPUT METHODS AND EQUIPMENT

Standard On-Line Printer
Datapoint Off-Line Printer
Computer Output Microfilm (COM)

The majority of all our output is on standard IBM 360 or 370 printers. Although much of our input is a mixture of UPPER CASE and lower case characters, most of the output is UPPER CASE only. At one plant in Copper Cliff, Ontario, Canada, we use a lower case print train. At other locations, lower case characters are printed as upper case characters.

Any of the three standard output formats (Master File, Register, Index) can be delivered on nine-track magnetic tape. The tape can be printed on the off-line, slow speed, typewriter style printer attached to the Datapoint Terminal in upper and lower case characters.

The tape may also be used on the Datapoint as source information to generate major blocks of new input information. On some large volume printouts, we send the tape to a service bureau and have it made into microfiche. We have standardized on a 42X magnification which produces thirteen rows and sixteen columns of full-size (15" x 11-1/2") computer pages per 6" x 4" fiche or a total of 208 pages plus a caption row at the top. These fiche are produced in cine mode and have both upper case and lower case characters.

INCREASED PRODUCTIVITY EQUALS REDUCED COSTS

Our Datapoint Terminal has allowed the greatest number of savings so far. Other manufacturers offer similar type equipment. It should be pointed out that this terminal is used for many other purposes than entry of SPINDEX data. This is in line with the current trend to multipurpose terminals most frequently expressed in "Office of the Future" concepts.

The following are the specific areas of input cost reduction and productivity increases:

1. Distributed Keyboarding
2. Typewriters as Keyboards
3. Sight Verification
4. Video Screen Corrections
5. Simplified Keyboard Input Forms
6. Elimination of Special Keyboarding Forms
7. Tutorial Data Entry
8. In-Line Batch Editing
9. Off-Line Batch Editing
10. Abbreviated Key Entry Supplemented by Machine Entry
11. Block Changes by Machine Entry

Further systems analysis will surely point the way to even greater improvements over the next few years.

Mariz: Thank you Dick. All right, questions?

Thompson: How much is the Datapoint 1100 per month?

Parmenter: If you have just the two-diskette with the keyboard and the communications attachment, which would allow you to go in batch mode, which is all you really have to have, that would be $400 per month. To add on the printer and the tape drive, which I have added to one of our units, it then goes to $800 a month. I mentioned, we're upgrading now to four diskettes and three additional terminals so that four people can input simultaneously. It goes to $1300, but then when you divide it by four terminals, it goes back down.

Max J. Evans: This basic Datapoint has a tape drive on it too? Is that right, or just a diskette?

Parmenter: The basic one just has the diskettes.

Evans: How do you convert the diskettes so that they're readable by the computer? What happens after you get it on the diskette?

Parmenter: How do you get it to tape, mag [magnetic] tape?

Evans: Yes, did you say you have a direct tie-in to the computer?

Parmenter: We can either go over the telephone lines in communications mode, or, which we do most commonly in our own building, we convert to tape and carry the tape physically upstairs and put it in the machine room. We have to write nine-track, 800 bpi [bits per inch] tape.

Evans: At what point does all of this interface with SPINDEX, and what program do you use? What front-end program is it? Did you have to write a new one, or is it one of the existing ARSP1 programs?

Parmenter: The card one [card loader program (ARSP1B)], we write on the tape as though it were a card image, eighty characters long, so that it's a minor modification to the JCL [job control language] to the card-read program, and the disk processor uses the same program. We still have it open so we can submit cards or tape and intermix batches. We submit four or five batches a day. We have thirty applications, and any given day we might be updating five different applications.

Frank G. Burke: One point, with the program at the NHPRC, we're using the CPT [model 4200 word processing typewriter], and we have two options there: just the straight CPT mag tape cassette typewriter or the same unit with a video unit, that displays fourteen lines at a time, attached to it.

Evans: What are the rental costs?

Dennis A. Burton: $375 a month. It's less if you buy it.

Burke: That's for the whole thing?

Burton: Yes.

Burke: With the video display, about $245 without it.

Burton: $250 without the video.

Evans: With that machine, after you record on the cassette, you need another machine, don't you, to . . .

Burton: Which the Archives [National Archives and Records Service] has.

Evans: . . . which the Archives has, but nobody else does.

Stephen Hannestad: However, it should be noted that CPT has, within the last six months, brought out a communicating system so that you can, again, go by phone line to the computer center without going through a media translation. So, that would require a minor modification to the ARSP1D program [CPT loader] because, right, 1D is looking for it to come in on tape, and if it comes in on the phone line instead, it's going to have to have some modifications made. But it's in the form of a minor modification, and, in fact, within the last several years, virtually every word processing machinery company has introduced the communicating option, and in most cases, they are now willing to communicate in either ASCII [American Standard Code for Information Interchange] or EBCDIC [expanded binary coded decimal interchange code], whereas previously, the communicating equipment required transmission in something called "communicating code." And "communicating code" was different, and if you had a Wang [Wang Laboratories, Inc.], and you had a CPT, they were not in agreement with each other. The current system, now, though, says that you might have a CPT that will talk to a CPT in its own language, but you can always flip the switch on the machine and use either ASCII or EBCDIC. I'm not saying that the machine is going to have all three, but a minimum, the communicating machines now have either ASCII or EBCDIC, so that the problems of going from a word processing typewriter into the computer are greatly reduced. And along the line, somewhere in the relatively high priority, we have determined that we are going to do the modifications necessary for CPT communicating typewriters.

Thompson: How much is the communicator per month?

Hannestad: The communications is an add-on box and, I think, adds $120 per month.

John F. Burns: It's worth $3,000 if you're going to buy it.

Hannestad: Right, in many cases, it is cheaper to buy than exercise the rent option. Now, we will be doing this conversion based on CPT, mainly because we have eight CPTs and zero anything else, and if we're going to pick a machine to write the programming for, selfishly, we're going to write it for what we've got eight of, and not for what we've got zero of. However, I think that the philosophy that we use in developing that interface will be transferable, even though the exact code structure will not be transferable from machine to machine. I have also found that if a man from Wang and a man from Raytheon [Raytheon Company] or whatever is trying to sell you his machine, on occasion, he will even tell you how to make his machine talk to your computer, because he figures he isn't going to make that sale until he tells 'em how to shake hands. We had the man from Xerox in to talk to us to sell us their fantastic new Xerox 800. And he left with his head between his tail or something because we tried three times to get his machine to talk to the computer. Three times it struck out. Before you buy or even commit to rent, force the vendor to prove that his typewriter can talk, not to a computer, but to your computer, because there is a big difference. When they were demonstrating their communicating capability, they were talking from an 800 to an 800 in the computer room. They were not talking from an 800 to a computer, and there's a real difference. Anybody can talk from typewriter to typewriter; that doesn't help you. But, by and large, the vendors now offer an adequate communications option. Most of them are willing to imitate standard IBM transmission characteristics. The compatibility is now beginning to exist. It wasn't there several years ago, and you had to go through the intermediate step of nine-track tape, but it's now available, and it's not hard to achieve.

Parmenter: I'd like to make one more comment. I'm trying, also, to acquire a portable terminal that I can plop down for a one-shot deal at a remote operation and get it going. You could enter outlying information and communicate in a batch mode to another terminal. I have two of them in mind, the Texas Instruments portable terminal . . .

Hannestad: Hopefully the bubble memory variety, I'd like to see how that works.

Parmenter: . . . and then the other one, something called "Computer Devices" [Computer Devices, Inc.], which has a little mini-cassette in the back of it with an acoustic coupler built in. So they can edit and do everything off-line. They can see what they're doing and revise it, and you can just pick this up in-case and flop it down on somebody's desk, when it's a brand-new job until they catch up. And then, when they've caught up with the volume, the work to update, to keep a data base running, is very small. But to get the original batch in is just a fantastic job. So these are two units, and

there must be more, I haven't investigated that closely, but I think they are worthwhile looking at to get over these humps. In a small remote group, at least, take it out to the field and allow them to enter their information out in the field and then just transmit it back over standard telephone lines to your central office. You'd be better off than bringing all this paperwork inside, trying to line it up and do the keypunching or editing or whatnot.

Hickerson: I think we had better move on to SPINDEX III.

SESSION IV

SPINDEX III

Ronald G. Watt: One of the reasons we came here is "What's next?" Our speaker, Stephen Hannestad, has already been introduced, and he's the man who can tell us about SPINDEX III. The question that we have had, and I'm sure that each of you has had is, "What will it do?" We have already had a slight introduction to it, so I'll turn the rest over to Steve.

Stephen Hannestad: "SPINDEX III"

SPINDEX III begins with the time-proven SPINDEX II data entry and files maintenance programs but offers much greater sophistication in both the register (text) listings and the index listings. Even the file maintenance portion of the system will be enhanced by the addition, in release 1, of a sub-file generator and, in release 2, by the addition of a selector/associator package.

SPINDEX II was designed to produce line-printer listings; that design concept severely limited the flexibilities in report formats. SPINDEX III is designed to utilize the vastly greater formatting and topography capabilities of electronic photocomposition for its listings. Both the register and indexing programs have been improved to provide greater flexibility and improved content. However, files produced under the SPINDEX II system will be fully compatible with the SPINDEX III system, and SPINDEX III files will be compatible with the existing SPINDEX II line printer programs (enhanced OS/OS-VS versions) when the sophistication of photocomposition is not required.

The first release of SPINDEX III is expected in July 1978. It will consist of the ARSP1A through ARSP4 input and files maintenance programs from the March 1978 release of the enhanced OS/OS-VS version of SPINDEX II and six new programs.

A. SPINDEX III Index Generator Program

SPINDEX II permits only single-level indexing, and entries under the keyword can be sorted only by title or date field. All fields in the master file are carried over into the generated index records. SPINDEX III will automatically generate dual-level indexing (primary keyword, thereunder by secondary keyword). Individual entries under the secondary keyword may be sorted by any designated field. The index sort field will consist of the keyword, secondary keyword, and tertiary sort field. The total length of the sort field has been expanded to 200 characters, and the length of each element can be varied from 2 characters to 99 characters. Index terms may be selectively generated from any SPINDEX level. Up to ten field types (first two digits of the tag) may be selected for generation of index terms, and the index function can be selected as follows:

(a) field permuted word-by-word

(b) entire field as one term (personal name field)

(c) field permuted for four-digit numeric words (year dates)

(d) entire field as primary keyword and secondary keyword, each element separated by a delimiter code

The generated main index term provides the primary sort element. The secondary sort element will either be the subordinate index term generated under index function (d) or blanks. The tertiary sort element consists of a parameter card selected data field.

Up to ten types of data fields can be selected for inclusion in the index record. These fields need not be the same as the fields used to generate the index terms. The entire input record is not carried over into the output to prevent the file from reaching an excessive size. To provide greater flexibility in designing index layouts, the first data field in the output record may be selected out of tag-number sequence. All subsequent data fields must remain in tag-number sequence.

The most significant improvement is the feature allowing two-level (main keyword/subordinate term) indexing via index function (d). In order to use this function, both the main keyword and the subordinate term are entered in the same field, separated by one of two special delimiter codes ($ or ¢).

Keyword$Subordinate term will generate two index cards:

Primary: Keyword

Secondary: Subordinate term

and

Primary: Subordinate term

Secondary: Keyword

This feature would be used when both terms are of equal intellectual importance and two index entries are desired.

When the secondary index term is meaningful only in the context of the primary index term, only a single index record would be created:

Keyword¢Secondary term will generate only one index record:

Primary: Keyword

Secondary: Secondary term

B. SPINDEX III Index Sort

This program will sort the index records on the 200 character sort field.

C. SPINDEX III Index Format and Photocomposition Interface Program

This program formats the index file, removing duplicate keywords, and within keywords, removes duplicate subordinate terms. The program permits the addition of a 600 character preamble to the print file. Both the keyword and the subordinate index term can be converted to upper/lower case. This feature must be used in conjunction with a special code to prevent undesired lower case characters (such as Roman numerals). The program can insert both typesetting codes and constant data (or headers) for each field in the file.

D. SPINDEX III Text Extractor Program

This program creates sub-master files by selectively extracting data from the master file. Using this feature, one can avoid processing the entire master file when only a portion of the data is required for a given project. Data may be selected by a combination of record level and field tag. For each record level selected, up to 25 types of data fields unique to that level can be selected. All data fields must be selected and maintained in tag-number sequence. In addition, the control field is automatically copied onto the extracted record. As the sub-file is in the standard SPINDEX master file format, it can be used as input to any of the processing streams.

E. SPINDEX III Text Format and Photocomposition Interface Program

This program formats the text file, removing file control data, and, under parameter card control, inserts file preambles (up to 600 characters) and typesetting codes. Up to 25 fields may be printed at each record level. For each field/record level up to 44 characters of typesetting codes and/or constant data can be inserted for the first occurrence of the field in a record (generally when the third digit of the tag = Ø), and 20 characters of typesetting codes and/or constant data can be inserted for subsequent occurrences of the field within the same record (generally when the third digit of the tag = 1,2,3 or 4).

F. SPINDEX III Extract by Control Number Range

This program will create a sub-file from a larger master file by copying all records from a given record to a second given record, inclusive. The first and last records in the sub-file are identified by the control fields for those records.

Both the SPINDEX III program F and program D can be used by SPINDEX II systems without modification. With some loss of format control and sophistication, the other SPINDEX III programs can be used by line-printer based projects by substituting printer control and no-op characters for the typesetter codes. The programs would also require modification for line-advance controls. For most applications, however, the SPINDEX II programs 5-8 will be the most effective for line-printer based systems.

SPINDEX III release 2

The second release of SPINDEX III is expected by July 1979 and will add one additional program, the selector/associator package.

This program will permit the selection (extraction) of records at a given hierarchical level based on the contents of a given tag. If the field contains alpha data, the selection routine will compare the selection key against the field's first N characters and will extract the record if an equal match occurs. If the field contains numeric data, the routine will permit equal, greater, or lesser comparisons of the selection key against the data field.

As the SPINDEX system is based on a hierarchical record format, the associated higher level records must be present in the extract file to provide the proper context for the records selected by the field comparison routine. Therefore, the associator package will select the associated higher level records. For example, if the selection module is used at level 5, the associator module will automatically pick up the associated level 1-4 records.

The associator module can also be used to link related lower hierarchical level records to a selected higher hierarchical level record. For example, if the selection module is used at level 1 (NHPRC institution level) to select records of a particular type of institution, all related level 2 records (NHPRC collection level) can also be automatically selected.

The input file can be either a full master file or a limited extracted master file.

John F. Burns: Talking about the ability to extract by control number range, can you do that more than once at a given time, for more than one range? In other words, say you've got numbers one to a hundred, just for illustrative purposes. Can you blot out three to five, fifteen to twenty, and sixty-three through seventy-four all at the same time?

Hannestad: Yes, no. (laughter) The answer is simply, for each single pass of the master file, you can get one range. There is no reason not to combine three extracted files back together again by just sending them through the same merge program. Now, use one of them as the old master, use one of them as the update master, create a new master, and come back in with the third file on top of it. Or if they're coming out in numerical sequence, you could just open one file and not put a file mark at the end of it. Bring the second file right in on top of it.

Burns: It can be done one way or the other?

Hannestad: Oh yeah. Yeah, the advantages of this, for example, you want to do a guide strictly to archival institutions. In your level one record you have a code which says, "I am an archival institution." It can compare against just the level one records and then automatically pick up the collections. At the level five or four, or wherever, you put in a subject term, "Apache Indians," or whatever, and if it matches on "Apache Indians," it will go back and pick up all the higher ordered records without which the entry just would not make any sense. On the other hand, if you're doing the extract on the holdings of a single institution, there would be no need to carry over the level one record. There will be considerable capability for saying, "Well, I want levels two, three, and four, but I don't want five, six, and seven, on whatever." Yes?

Margaret Saunders: How many fields will you be able to print?

Hannestad: You'll be able to print twenty-five field types at any given level. Now, if you are making a distinction between 350 and 35-anything, then you will have burned up two of your twenty-five at that record level. If you want to take 35A, 350, and 35-anything else and do three different things, then you will have burned up three of your twenty-five at that level.

Joan M. Winters: Does that mean you can specify all three characters of a tag?

Hannestad: Yes, however, the price is: anytime you specify to three digits, you have specified the unique field, and you have burnt up one of your twenty-five tries at specifications. So, I would not recommend that, unless you had some real purpose to call it out, you would use it. But, at twenty-five fields per record level, you can see that there is now a tremendous amount more that goes into a print file than previously went to a print file.

Watt: I have a question on costs. If you buy SPINDEX III, does that mean both the first release and the second release . . ..

Hannestad: Yes.

Watt: . . . or is there additional cost?

Hannestad: No. The second release is considered to be part of the SPINDEX III system, and again, as far as the pricing policy goes, it looks, now . . . Frank, do you have the cost tables?

Frank G. Burke: Nope, I don't have anything. (laughter) I'm just a customer.

Hannestad: . . . ahh, first of all, anyone who has SPINDEX II will be given full credit against the cost of SPINDEX III. Somebody who buys SPINDEX III, who does not have SPINDEX II, is going to get SPINDEX II as part of the package anyway.

Burke: Aren't we going to deduct "wear and tear?" (laughter)

Hannestad: No. The same pricing discrimination will exist. Non-profit institutions will get a discount; for-profit institutions will pay full price. I believe that the total SPINDEX III price is in the realm of $1500.

Brent G. Thompson: For non-profit?

Hannestad: For non-profit the price is $1500. Now, in addition, our policy is that SPINDEX III will be sold only with training bundled as part of the price. If you buy SPINDEX, you are entitled to send two people to the training course in Washington [D.C.], and it will be a one-week training

course. Now, if you don't send the two people, we don't give you a discount. But what we're saying is, by bundling training in, this give us a mechanism to set up a formal training program. One of the, probably the biggest deficiency in SPINDEX II is not in the code; it's not in the documentation; it's in the fact that it was released without any form of training with it. This is one of the mistakes that we're going to correct with SPINDEX III. Training will be available ahead of time for SPINDEX III.

Burke: Say something about documentation. (laughter)

Hannestad: Yes, I believe that's tomorrow morning's subject, Frank.

Arthur D. Dick: What kind are you going to be offering? Are you going to be offering the "I want to know how to code" training, or the programmer that's going to have to run it training?

Hannestad: No, the programmer that's going to have to run it is, well, as Dick has pointed out, the programmer plays his games with OS (Operating System) and bows low to the computer in the morning. Forget him. The person who needs the help is the one who's going to have to say, "Gee, how am I going to set up my control number structures? How am I going to use my tag number structures? How do you go about using electronic photocomposition? If you want to use italics, how do you go about doing it? How do you design backwards from a typography that allows you, from '18 pt. bold' centered, over two columns, down to '6 pt. italic' with everything in between?" To make that conversion, from a typography design back into a data base that is going to work, is going to take some training, and it's not going to be possible just to try to write a manual that explains it all.

Watt: How long do you expect the training to be?

Hannestad: We expect it to be a five-day training session.

Thompson: When will it be offered?

Hannestad: We're trying for July, are we not?

Burke: Yes, there is the point Steve, the point that we were discussing last night about the possibility that instead of people coming to Washington for

training, if there was enough call for them, that it might be that the training would go to the regions.

Hannestad: Yes.

Thompson: How many times a year would it be offered?

Burke: It depends on the number of sales of SPINDEX III.

Hannestad: I think it's safe to say that it's highly impractical to hold a training session when you have less than ten people to be trained. Something in the realm of ten to twenty-five people being trained at one time would be optimal. So it's a question of arranging; when there's a demand it will be met.

Dick: For whom is this training intended? Will it be suitable for people who have never touched a computer or only for people who are already familiar with SPINDEX II?

Hannestad: No, it might be a little repetitious for people that are associated with SPINDEX II, but it's going to be designed from the idea of: "What is a control number; how does it relate to a manual outline?" Right, back to basics. And you know, frankly, the way the parameter cards can code themselves by answering an English-language question, the actual instruction needed to code a parameter card is not that much. The instruction that's needed is in how to get from the concept of typography and data base design and what it really is you want the system to hold and how you want it to come back out that's important.

Bruce W. Dearstyne: Do you have any idea how many of these are going to sell? Do you expect them to sell like hotcakes? (laughter)

Hannestad: I don't honestly know what the sales record will be on them. I think a lot of that depends on how well a national data base catches on; how many regional projects get going that are actually going toward the complexities of electronic photocomposition. Continuing to use just the SPINDEX II routines, there will be advantages to be gained from SPINDEX III, even on the line-printer by being able to refine your data base, make your extracts, and make your selections.

Burke: I'd like to make that a little clearer. We never got into this thing to come up with a better product because the one we have isn't selling very well. We got into this because, even when SPINDEX II was originally being designed, there were certain things that we wanted in it that never got into it. We weren't talking about photocomposition at that time, but we were talking about the selectivity and certain indexing features and the two-level index. We just never got around to doing those things. As the need showed up more recently to do them, what we wanted to do was to improve the program, essentially again for our own purposes, but, being a government agency, we should make it available to those who need it on an as-needs basis and at as low a cost as possible. There's some cost because we have costs involved. So that's the way it came about, and then we just said, "Well, on an as-need, we'll make it available to anyone who wants it, but it would be unfair just to dump it in their lap, so we'll also make available with the cost the training for using this thing." We might not sell any for the next ten years. I don't know.

Hannestad: It all depends upon how our salesman does. One other thing that should be pointed out is: with SPINDEX II we had several project-oriented, special interface programs for the Government Printing Office's Linotron 1010 made. I don't know how many Linotron 1010's there are, but I think there are three, and two or three of the three are in government captivity. When we went into this system's design, one of the first things that we did was to make certain that there were commercial, outside vendors with available equipment that it would be compatible with. We were content when we found three bidders who all said that they could handle it. We're not necessarily saying that these three companies are the three companies that you should use, but we have definitely established that there is a commercially available way of getting electronic photocomposition done which does not rely upon using government equipment. We are planning on using one of these vendors instead of the Linotron because we feel that we can get better service.

Richard S. Parmenter: Is it a secret who the three vendors are?

Burke: We have a contract, or we're about to write a contract to do the guide.

Dennis A. Burton: Our contract will be with International Computer Print Corporation in Fort Washington, Pennsylvania.

Parmenter: Are they using their own equipment or somebody else's?

Hannestad: Their own equipment.

Burton: VideoComp 500 Photocompositor [RCA Corporation]. I can't give you any other names; that's the one I know. That's the contract we have, but there have got to be lots of other folk out there with VideoComp 500's.

H. Thomas Hickerson: You say three vendors, are you actually talking about three commercial firms or are you talking about three lines of equipment?

Hannestad: We found that we are totally compatible with Astrocom [Astrocom Corporation], VideoComp or Photron. So it's not just that there are commercial service bureaus that will do this work, but, by using the capability to insert whatever typography codes are needed for a particular machine, you can give them an "ASCII bell" if they want an ASCII bell." If somebody else is willing to accept a "@" as the "here's the beginning of my code" sign, you use that. Because it's in the parameter cards, not in the firmware or the software, the interface becomes totally independent of the machine it's set upon. Now, in our practice, the vendor said, "Why change? You're using a particular set of typography codes today; we're compatible with you." If we were bringing up a totally new project, we might have used his native coding structure rather than the one we had been using for the last three years.

Burke: So the vendor accommodated us instead of our accommodating the vendor.

Hannestad: The point is that this is not a system that requires you to use only government-owned equipment.

Burton: For those of you who might be interested in a price, the per-page composition in our sample Directory of Archives and Manuscript Repositories, either for the two-column, textual portion or the three-column index, is $4.70 a page.

Hannestad: On the other hand, the last time we had something set in "cold type," what did we pay, $35.00 a page?

Burke: We don't set anything in "cold type" anymore.

Hannestad: The point of the matter is: not only is it cheaper, but because it is never rekeyboarded, you do not have to edit for the introduction of new errors.

Burton: Another advantage is time. Our publication will be about 900 pages long. We will have the camera-ready copy ready to go on the presses three to five days after we have completed our processing, rather than waiting weeks or months for the composition.

Max J. Evans: Is that guaranteed in the contract though?

Burton: Guaranteed in the contract.

Hannestad: Complete with penalty clause. The other thing is that, just because the page has been photocomposed, there is no reason that a paper product ever has to result from it. You can use that as the input to microfiche or microfilm. At $4.00 and something, it's cheaper than typing it.

Hickerson: We must leave before they lock the doors. Tonight's dinner is in the West Lounge of the Statler. From the main desk, it's down the hall to your right, and the bar opens at 7:00.*

*During the "cocktail hour," Richard Parmenter provided a slide presentation illustrating implementation by INCO of the SPINDEX II program package. Samples of selected applications are included in Appendix C.

SESSION V

Technical Aspects of SPINDEX Use

[In this session, chaired by Joan M. Winters, there were no formal presentations. The following papers had been distributed and were discussed during this session.]

Joan M. Winters: "Cornell's Experience with SPINDEX II: Problems and Extensions"

The Office of Computer Services (OCS) at Cornell University, under the auspices of the Department of Manuscripts and University Archives, has developed two applications of the SPINDEX II program package. In addition, OCS has performed varied testing of the programs, most recently on the GS3005 version (distributed to us on June 21, 1977).

Testing of this new version has shown that some problems occurring in an earlier version (NASM08)[1] have been corrected. Specifically:

> In the ARSP8RSP program the number of lines printed per page can be increased from the default of 55 without causing a spurious syntax error.
>
> It is now possible to print in position 132 without the ARSP8RSP program's giving a spurious syntax error.

Testing also indicated problems in the present version of SPINDEX II. These are summarized below. Some of the problems can be resolved by updated, expanded, and reorganized documentation. Others can be reduced by improved error diagnostics. Still others result from limitations in the design of the programs. Finally, some are errors in program execution which must be corrected by code changes, an approach that Cornell has not yet taken-- believing it more efficient to stay with the standard package. Some program errors are just a nuisance. Others are more serious and limit the type of output that can be produced.

The problems outlined below are arranged by program:

ARSP1BSP

1. This program abended with an OC5 completion code when a tag contained more than 993 characters of information. A more diagnostic error message is desirable.

[1]H. Thomas Hickerson, Joan Winters, and Venetia Beale, SPINDEX II at Cornell University and a Review of Archival Automation in the United States (Ithaca, New York: Department of Manuscripts and University Archives, Cornell University Libraries, 1976).

2. It would be helpful to indicate in the documentation[2] that blanks entered at the beginning of a tag's information field are retained, although only one blank is kept if several are entered at the end of the line of a tag which is continued.

ARSP2RSP

1. A record which apparently contained only 6996 characters was rejected as being too long. A record which apparently contained 6981 characters was accepted.

2. Not all situations which are potential problems are diagnosed. For example, when automatic generation of tag 400 has been selected and a tag 400 exists in the input stream, no warning message is issued.

3. The fact that the tag site number can be alphabetic is not documented, nor is the fact that such an alphabetic character must be capitalized. (This constraint becomes a problem with interactive data entry.)

ARSP4SP

1. The SCCS code is not always printed, e.g. for tags 120, 630, and 640, although the code is printed for tag 110. This problem made proofreading the data difficult.

2. The documentation on the meaning of the hierarchy levels and their integral relationship to the control number should be expanded. In particular, I was unclear about whether the hierarchy level, in conjunction with the control number, could be used to identify a record uniquely. It cannot. Nor did I realize that the same tag cannot be used in the same record at two different hierarchy levels. (The second overwrites the first with no error message issued.) As a result of these limitations, I could not implement a proposed SPINDEX application, at least according to my initial approach. As a final point, no indication of hierarchy levels is printed by ARSP4SP, except for the first encountered in a record.

3. When a record was truncated, not only was the last part of the information lost; but the first four lines were also apparently deleted from the last tag.

[2]The SPINDEX documentation referred to here and subsequently is National Archives and Records Service, SPINDEX II: Report and Systems Documentation (Washington: National Archives and Records Service, 1975.)

ARSP5RSP

1. When the program was instructed to print in position 132 via the AR5D and AR5E control cards, the program consistently terminated with the error message "Control fields in card not all numeric and*or modified length of field exceeds 132 print positions."

2. If the number of lines printed per page was increased above the default, e.g. to 57, the program terminated with the same error message. However, the program did successfully print a number of lines per page which was less than the default, e.g. 50.

3. The level 5 position of the control number did not print when it was requested on the AR5D card.

4. Although tag 100 information is indented according to the hierarchy level, the tag 910 series information is consistently indented only two positions from the position requested. Thus if a register contains records at different hierarchies, the printed output looks visually inconsistent. These features caused a problem in Cornell's Day Papers application, which was resolved by editing the register produced by SPINDEX II interactively to produce the published register.

5. If a record was diagnosed as having "exceeded maximum length" by ARSP2RSP, ARSP5RSP abended with an OC7 completion code when it got to the end of the truncated record.

6. When a tag contained no blanks in its information, the program printed one character per line rather than generating an error message.

7. It would be desirable to indicate in the documentation that a tag with an alphabetic site number prints before the same tag with a numeric site number. Otherwise people using the multiple abstract feature may be surprised at the order in which the information prints.

8. The default print position for printing dates starts in position 76, not in position 86 as indicated in the manual. Likewise, the default print position for the header DATES is position 78.

ARSP6SP

1. The program does not generate keywords from additional tags, contrary to the documentation. This is a serious program error which could make infeasible some applications not very different from those already developed at Cornell.

2. It is most desirable that the requirement of entering tags in ascending order be clearly documented and emphasized. (I have only found the restriction mentioned in Appendix III of the manual for ARSP5.) Prefer-

able would be removal of the requirement entirely, for increased flexibility.

3. When a tag permutes word-by-word, punctuation such as commas and apostrophes cause a new keyword to be started. Since the Cornell applications use word-by-word permuting, this feature has required entry of personal names in a non-standard format, e.g. O'Hare, William becomes O_Hare_ William. It would be desirable for these characters to be considered part of the keyword.

4. Keywords were generated from tag 210 when the permuting function was changed on the AR6B card, even though tag 210 keywords were not requested on the AR6A card.

ARSP8RSP

1. The last letter of the first line of the title is truncated if the last letter of a word falls in the position equal to the length of the title field plus 1, e.g. in the forty-third position for a Cornell application. Cornell must correct (or get a correction for) this bug before the application can be successful.

2. When the print positions for series, box, folder, or repository were changed from their defaults, they each printed with a trailing garbage character. (The problem can be bypassed by specifying the lengths as one less than the actual lengths.)

3. The fact that tags must be entered in ascending order is not documented. It would be useful if this restriction were removed.

4. Coding an X in column 8 of the AR8T card did not completely suppress the index. The keywords were still printed, overwriting the first line of the entry below the keyword. The rest of the index was suppressed.

5. The abstract (tag 900) did not print when printing of a special field was requested on the AR8A card but no particular special field was specified on the AR8D card. No warning message was issued for the missing tag on AR8D.

6. Except for the first line, the length of the printed title field is one more than the length requested. Thus the title field can extend three positions to the right of the right-most position intended and therefore unexpectedly overlap the next field. This implementation differs from that of other tags, e.g. tags 250 and 630, which contain only the requested number of characters in the printed index. If these tags are actually longer than the printfield, the additional information is truncated without a warning message.

7. The tape copy of the formatted index, produced through DDcard AR0016,

did not contain title and associated information for those records which had an abstract (tag 900). The abstract had been specified for the index and had printed successfully in the printed index. Information for an additional record was also missing.

8. The number of lines printed per page was generally four more than the number requested. (This was not true for ARSP5RSP.)

9. The default headings for dates and collection identification were DATE_ and COLL ID.

10. It would be useful if all of the fields which are standardly printed could be retitled, e.g. LVL, QTY, and REPOS. It would also be useful if fields could be titled with more than eight characters.

11. It could be useful if the abstract (tag 900) could be printed under the title (tag 100), as it is in ARSP5RSP.

Another problem Cornell has encountered is not associated with any individual program. When new versions of the package are distributed, it would be very desirable to have more supporting information sent with the tape, such as the version number, a complete description of enhancements and how to use them, and the function of each routine on the tape. This procedure would allow people to take advantage of new features and save time.

In addition to the problems noted above, Cornell has found it necessary to extend the SPINDEX II package for input and output. I have done so without modifying the programs, by taking advantage of the interactive computing systems available at Cornell. Originally IBM's Time Sharing Option (TSO) was used to edit the guide to the Day Papers produced by ARSP5, before publication.

More recently, I have used the interactive system now available at Cornell, the Conversational Monitor System (CMS), to create entire SPINDEX II job streams and to modify the index printout. (CMS runs under the IBM operating system (OS) called Virtual Machine (VM). Cornell also runs OS under VM.) In particular, I have used the EDIT facility of CMS to create data for input to ARSP1BSP that look as if they are on cards. The procedure is awkward but feasible. Taking advantage of some local EDIT macros, I have also used EDIT to modify the Job Control Language and SPINDEX II control cards necessary for a SPINDEX job. When a job stream is complete, I can submit the job directly from the interactive system to OS. When desirable, the printed type of output can be returned to the interactive session. In an application currently under development, it is necessary to do this to modify some of the titles in the index. Special EDIT macros were written to simplify the process.

Additional CMS facilities could be developed relatively easily to simplify and tailor these extensions further. As time and money allow, some enhancements will probably be implemented.

Richard S. Parmenter: "Modifications Made by INCO for Converting SPINDEX to RIM/DEX"

List of Changes Made to SPINDEX

1. The control number was changed to accept "-" (dash), "." (period), "?" (question mark), as valid characters.

2. In the Number 8 program, we substituted the literal "LOCATION" in place of the literal "REPOS".

3. The original DOS version allowed imbedded blanks within the control number. The OS version did not. The OS version was adjusted to allow imbedded blanks in the control number.

4. We changed the position in both the Number 5 and the Number 8 programs of date and page number. Date was moved to the left-hand side, page number was moved to the left, far enough to allow it to appear on the right-hand side of a computer printout which is trimmed to 8-1/2 x 11" page size.

5. An apostrophe ("'") was changed to be a valid character when the number 6 run was generating keywords. For instance, this would allow O'Neill to be listed under the O's instead of the N's. We experimented with "." and "/" to be valid characters in the selecting of keywords. We later reversed our decision and decided not to make these valid characters for keyword selection.

6. We duplicated the multiple 900 tag routine from the Number 5 run to the Number 8 run. This means that in the Number 8 run, the 900 tag starts on the next line after completion of the 100 tag and prints multiple 90X tags.

7. We established the "." in the first position of Control #1 when it is a new keyword from the previous record to cause the print to jump to a new page. We call this our page break routine.

8. We have established a sort card in the Number 7 run to be submitted with every job just as though it were an AR7 card.

9. A new parameter card, AR8F, was established which allows a title to be entered on the Number 8 run. Formally, every Number 8 run had a constant title, namely, "INDEX".

10. The "_" (Hex 6D) was blocked from being converted to a (Hex 38) on the number 2 run. This overcomes the problem of suppressing an underscore character from printing if used in a tag greater than 200.

11. The position of the stopword frequency count and the asterisk used to denote stopwords were moved to the left-hand side of the page ahead of the actual keywords. This is a standard procedure to reduce parallax errors when reading.

12. In the OS version, it is possible, through the AR6A parameter card, to place either the 1XX tag or the 250 tag into the second keyword position. We have converted this so that any tag may be placed in the second keyword position through the use of the same AR6A card.

13. Originally, the number 300 tag (QTY) had an unusual print routine. This was eliminated to make the 300 tag similar to all other surrounding tags.

14. We are currently testing a new routine which would allow tags on the AR8D parameter card to be out of ascending order. This last change has not been made operational at this date but has undergone a number of successful tests.

Additions to the SPINDEX Program

We have added a new module which has an AR9 series of parameter cards. This new module allows one of three separate separations according to the way the parameter cards are completed.

1. The deletion of a specified tag from all items in the data base which contain that tag.

2. A global change to the variable data portion of all records in the data base by substituting a specified new string of characters for a specified old string of characters wherever they may occur.

3. The selection of a single predetermined contiguous range of control numbers from the master record which can then be processed by any of the other standard modules of the SPINDEX program such as the Number 5 or 8 runs.

Supplemental Technique Using the Datapoint Terminal

We are able to take magnetic tape versions of the Number 5 or Number 8 runs; and then using the stand alone program, generate deletions, changes or additions of tags for subsequent entry into the master file. This allows major block changes to be done by machine rather than corrected individually by the keyboard operator.

List of Potential Future Changes to RIM/DEX

1. A vertical print format which would: a) allow any selected tags, b) in any tag order, and c) allow all tags to be multilined. Although this is number one priority in terms of potential improvements, the major nature of this change may well be put off for additional years unless unexpected support is received.

2. A change in the Number 8 program which would allow the AR8C parameter card to operate to allow header suppression when the header length is made equal to 00. This routine is already incorporated in the AR5E program and could be transferred.

3. The body of both the Number 5 and the Number 8 runs suppress the printing of the underscore character. However, underscores in the headers of both the AR5 and the AR8 print the underscore character instead of suppressing it. This should be corrected.

4. A position on the AR8A parameter card should be assigned so that the page break is on or suppressed depending on whether it is blank or marked with an X.

5. The hierarchical levels 0,7,8, and 9, should be able to be activated or disabled on the AR6A parameter card. They are now in a constantly activated condition.

6. Provision could be made to position or suppress the keyword in the Number 8 run.

7. The job stream should be established to allow a sorted output of the Number 7 run to feed into the Number 5 run.

8. Provision should be made in the Number 5 run to print and position the repository and the Collection ID codes.

9. The ARSP4 master printout should be revised so that tags with multiple lines break between words as they do with 1XX and 90X tags when printed in the Number 5 and the Number 8 runs. Currently, lines break at a fixed length which is frequently in the middle of a word.

10. Our own Number 9 modules run needs to be changed to accept several contiguous ranges rather than a single range. It also needs to be adjusted for the alternative of selecting everything except the range or ranges specified.

Dennis A. Burton: "Processing Problems: NHPRC Guide"

The NHPRC guide project will require major modifications to several SPINDEX

II programs as well as the creation of some new program units. These modifications and additions will be presented as a description of SPINDEX III.

Although our data base was designed for SPINDEX III, all processing to date has been done with SPINDEX II. The project does not use all of the programs. Data is loaded with ARSP1D; ARSP5 is never used. Additionally, two special programs select data from the file and code it for high-speed photocomposition.

We have encountered several problems in processing data through SPINDEX II. Most of these problems were caused by our own errors. To overcome the problems we in some cases altered input procedures, but in others we asked for and received program modifications to soften the impact of the errors.

1. Field Length Minimum

   In order to save keystrokes, single character codes were used in some select fields. These fields disappeared from the file during processing. We later learned that there must be at least two data characters in each field.

2. ARSP8 - Truncated Words in the "Title" Field

   When the title field exceeds the number of positions allotted it, additional lines are printed. There is a routine to avoid splitting words (i.e., the line does not necessarily go to the last position). However, when a long word fell at the end of a line we lost a character in processing.

   Our solution (required by the composition program) was to increase the title field to 99 positions and print all data on a single line.

   The program was modified to eliminate the truncation of words in ARSP8.

3. Index Secondary Sort

   We thought we could use the tag reassignment features in ARSP6 and ARSP8 to generate index listings other than alphabetical by "title" or chronological by "date". We found, however, that there are only two secondary sort options--by the lowest 1xx tag or by tag 250.

   For the present we live with the "title" sort (we do not use tag 250). For the future we have requested modifications to permit free choice of any tag for an alpha or numeric secondary sort.

4. First Character Capitalization

   There is a sub-routine in ARSP2 to raise the first alpha character in

some fields to upper case. Because this feature created unnecessary inflexibility, we requested its removal.

5. Special Characters

For our final product some characters were not readily available--open quote, close quote, and a non-printing character.

ARSP1D was modified to convert lower case key 41 to EBCDIC "4C" and upper case key 41 to EBCDIC "6E". These provide the open and close quotes. When using print train "TN" they appear as ( ) and ( ).

For a non-printing character we use upper case key Ø which appears as ( ) on print train "TN". We use the key to extend one-character fields or to alter numeric sorts.

6. ARSP2 Error Messages

We found that error messages were in some cases not printing in the ARSP2 diagnostic and error listing, and in other cases were incomplete.

The program was modified so that all error messages are now fully listed.

7. Field Exceeds 985 Characters (ARSP1D)

Exceeding the character limits of a field produced results that created serious problems for us:

a) Since no error message was printed, we could find these errors only by proofreading the processed material.
b) When the character limit for a field was exceeded, that record was truncated, dropping all remaining fields.
c) The truncation of the record meant that the end-of-record code was not read. Thus, two records were combined with the first of identical fields bypassed.

Since the truncation took place in APSP1D for which there is no error listing, programs 1D and 2 were both modified to soften the effects of this type of error. The programs will now accept the first 985 characters of a field and bypass the rest. However, the next field is read and an error message is printed in the ARSP2 diagnostic listing.

8. Wrong Length Tag in Last Field of Record

When a wrong length tag appeared in the last physical field of a record, the field was bypassed, but so was the end-of-record code, resulting in combined records. No error message appeared.

The program was modified so that only the field was bypassed; the end-of-record code is read. An error message is printed in the ARSP2 listing.

We expect to continue to find processing problems. We would like to think that our staff is creative enough to come up with errors and/or applications that were not anticipated a decade ago. At the same time, however, we expect the response to our programming problems by GSA programmers will be as quick and effective as it has been to date.

Joel A. Shirley: "Program Additions, Modifications, and Problems at the South Carolina Archives"

1. The first program modification we made changed the standard Spindex II control number subfield allocations from 4, 4, 4, 3, 3, to 4, 3, 4, 5, 2. Parameter card changes were then made for print programs ARSP5 (Register Listing) and ARSP8 (lndexer). We did this because more of our records would "fit" into the 4, 3, 4, 5, 2 format than the standard format.

2. As a result of this change and the fact that parameter card changes alone would not allow printing of the output data (contrary to what the systems manual states),[1] we had to modify the program ARSP8 sub-routine that moves the series ID to the print area into a sub-routine that moves all 18 characters of the control number. ARSP5 was also modified accordingly.

3. We print all of our 100 tags (description) in lines of 50 characters each. For some reason, if the 51st position of the tag is followed by a space, the 51st position disappears in the ARSP8 Indexer program. Fourteen instructions were necessary to correct this problem.

4. Another problem also involves our print format but has not been corrected. If a 100 tag line being printed ends on a multiple of 50 (i.e. 50, 100, 150, 200, etc. positions), a blank line is "printed" after the last actual print line. This takes place both in the ARSP5 and the ARSP8 print programs. We decided that at least for the present this problem would be more difficult and costly to correct than it would be worth.

5. We dropped program ARSP7 (keyword sort) and added our own sort package which takes about one-third of the sort time.

6. We discovered that if we encoded indexing information from two (2) entries with the same control field, they were merged into one garbled entry. We wrote a program to check for duplicate records that is used whenever this problem is suspected.

7. Since we wanted our output in the form of 16mm computer output microfilm,

[1]National Archives and Records Service, Spindex II; Report and Systems Documentation (Washington: National Archives and Records Service, 1975), p. 10.

a program was added that builds a Register Microfilm tape and an Index Microfilm tape in lieu of or at the same time as the print formatted tape is being created. This program utilizes the CM-360 Microfilm Management System.

8. We had a program written that lists input data in order for it to be proofread.

9. A tape management system was added to enable us to keep track of created master files (both completed and not completed). Included in this management system are different data set names for each job and a grandfather, father, son tape protection system.

10. We recently modified program ARSP5 to allow for Register listing of more than 999 pages. (The fact that only 999 pages were allowed for the listing was not noted in the systems manual.)

Brent G. Thompson: "SPINDEX II Program Modification"

The L.D.S. Church Historical Department has made two major modifications to the SPINDEX II programs. The first change adds a blank line between records in ARSP5. The separation between entries makes the scanning or reading of each page easier and improves the appearance of the output. This modification does not work when the control number is printed limiting its potential usefulness.

The second modification is much more important to the Historical Department. When we began testing SPINDEX II for use as a card catalog replacement, several problems were found with the pure EBCDIC sort of keywords. We found that SMITH, JOSEPH FIELDING, 1838-1913 appeared before SMITH, JOSEPH, 1901-1944 in the index. This occurred because the space between Joseph and Fielding in SMITH, JOSEPH FIELDING was being compared with the comma after Joseph in SMITH, JOSEPH. We knew that we would have over one hundred entries for the two men in the index, plus we had other similar cases. Other sorting problems included NAUVOO LEGION appearing before NAUVOO, ILLINOIS and WEST JORDAN, UTAH sorting above WEST, HENRY, 1903- . In each case the use of punctuation caused the problem. Therefore, we could either not use punctuation or modify the programs. We felt that punctuation increased the readability of the index so we decided to modify the programs.

We discussed our problems with a programmer who suggested that the punctuation in each record could be changed before the keyword sort and then the original punctuation returned before the index was printed. We found that if the commas in personal names were replaced by spaces that SMITH, JOSEPH, 1801-1844 would appear before SMITH, JOSEPH FIELDING, 1838-1918 and other personal name sort problems would be eliminated as well. We already knew that if the comma in NAUVOO, ILLINOIS were removed, then NAUVOO, ILLINOIS could appear before NAUVOO LEGION. We changed our tag structure so

personal names would be placed in the 100, 120, and 210 tags with place names appearing in the 200 tags.

The programmer added several sections of program code to ARSP6. In order to maintain the integrity of the original program the new sections are self-contained. A new sort key of a fixed 150 character length is added to the front of each existing record. The original record is moved into the new sort key with blank spaces replacing commas for the 100, 120, and 210 tags and commas are removed from the 100 and 200 tags. The integrity of the original is preserved. The records are then sorted in ARSP7 but before proceeding to ARSP8 the program exit E35 is used at which time the 150 characters added to the front of each record is deleted. The original records are printed in ARSP8 in the order the modified records were sorted.

The modifying of the SPINDEX II programs was more difficult, expensive, and time consuming than we had expected, but we are pleased with the results.

SESSION VI

The Future

H. Thomas Hickerson: Steve Hannestad is going to speak briefly about the role of NARS in the future maintenance and development of SPINDEX II and III. I think that almost all of us, at one point or another, have been critical of NARS' role in the maintenance and development of SPINDEX. At the same time, SPINDEX does exist, does work, and people are doing impressive work with it at relatively little cost. It is in hand, and it is getting better all of the time. So, while I think that criticism is healthy and that effective maintenance and development of any computer system are very important, NARS' role in this process has improved significantly. I would like to express my appreciation to all those people who have been involved in providing us with SPINDEX. Steve.

Stephen Hannestad: "The Role of the National Archives in the Future Development and Maintenance of SPINDEX"

I think that there is no way that I can describe what NARS' role in future SPINDEX maintenance is going to be without digressing and providing a little bit of background information. Contrary to some opinions, the National Archives is not a software firm. We at the National Archives do feel a certain responsibility to experiment and innovate in many areas of archival technology. Computerization is one of those fields in which we are doing some innovation. We are also attempting to do innovation in other areas, document restoration, document preservation, document reproduction, and we certainly wish to share the results of our experimentation as widely as possible. However, we are not set up in such a way that we can respond to an external request for system development. It has to be somewhere in an area that we have a current operational or research interest within our own requirements.

SPINDEX was initially conceived with the intent of solving some of our requirements for automation. We knew that it would never solve all of our automation problems. It was also originally conceived as a system that would be usable, not only by the National Archives, but by a wide variety of other institutions. Frankly, at the time we developed it, no one expected it would find any commercial application whatsoever. We are amazed by the things that International Nickel does with the system. But, as I say, SPINDEX was originally designed around archival problems and specifically with the intent of producing indexed finding aids to archival or manuscript holdings. Whatever else can be done with SPINDEX: "beautiful," but its central design concept was limited to guides, indexes, calendars of documents, chronological listings, etc.

As we grew with SPINDEX, we found various things in the original version that did not meet our current requirements. For those of you who have been with SPINDEX almost from the beginning, you will know that we have made the change from a DOS [disk operating system] version to an OS version, and

that we went from a thirty-field limit to a fifty-field limit. We went from a roughly 2,000-character record to a 7,000-character record. We expanded the size of the index sort field. Many changes have occurred over the years. These changes have predominately been as a result of applications at the National Archives. The keyword field was expanded when we started trying to index German, bureaucratic titles which could run on for pages. The fifty-field expansion came up when the Papers of the Continental Congress needed extra space. The 7,000-character expansion came along when the "Germans" decided that they needed to describe the bureaucratic evolution of an office and just could not live with only 2,000 characters. They said that the names alone took up half of that. The changes that we have made in SPINDEX have been generated by our own current applications' needs.

SPINDEX III derives from the same philosophy. SPINDEX III was designed specifically to meet the needs of the Commission's guide project, but, at the same time, we made serious efforts to include features that we thought would be usable or needed by others, but that we did not necessarily need ourselves at the time. By improved communication, by our being more familiar with what you are doing, we can take into account what your desires are for additional modifications and can consider whether or not they have direct application within the National Archives. Unfortunately, there are many things which I can say would be nice features for SPINDEX to have, but, if they have no application within the National Archives, we are prohibited by law from making those changes. The Office of Management and Budget is very much concerned that the government does not get into the position of competing with private industry, and so they put certain limitations on us that prevent us from providing some of the modifications that you have identified as being desirable.

SPINDEX II, as of this date, is supported by the National Archives in only one version. The March 28, 1973 version is the version that we are currently supporting. [Stephen Hannestad brought copies of the latest system tape to Ithaca. This corrected version was made available to users for them to copy.] Any bugs or problems that are found in this March 28th version will be corrected; the updates will be provided to all current SPINDEX users, and then that version, of whatever date it is, will become the one and only SPINDEX version that is supported by the National Archives. We cannot be in a situation where there are eight, nine, ten, however many different generations of SPINDEX and try to support them all. We can only support the most current version that we are, ourselves, using. Consequently, to any users who have made internal modifications to SPINDEX, we are not able to support your version of the program. We cannot provide maintenance for it. To the extent that a problem that we find in our version is likely to be a problem that you will have in your version and that the code that corrects it in our version will correct it in your version, we will disseminate information concerning what updates and changes we make. But we cannot debug a problem if it exists in your version and not in ours. Consequently, if anyone finds a problem in their SPINDEX system, what we are requesting is that the problem be brought to our attention; we will test it against our current version of the system; if, in fact, we find it to be a problem in our current version of the system, we will make repairs. We will notify everyone who has a version of the system

as to what the repairs were and how they were accomplished. If, however, when we test the problem against our version of the system, the error does not occur, our response will be that we are sorry, but our version handles it. Then you will be in a position where you will have to find a programmer to make the change in your version of the system or change to using our version of the system.

I wish that we were in a position to offer more support, but, unfortunately, we are not established as a systems design group. I think, right now, within the National Archives, I am the entire systems shop for systems design, and approximately 40% of my time is taken up on other administrative matters that have nothing to do with archival automation. Within the Office of Data Systems we have excellent programmers; however, we also have other applications at the National Archives. We are developing a new system for the National Audio-Visual Center. We have a system that has been developed for the regional records centers. We have many different demands on our programmers' time other than the SPINDEX system. So there is a certain order of priority. We will definitely do our utmost to respond to you, to help you in setting up SPINDEX applications, but I can give you no promises that, if you tell me today of a bug in the program, I can give you a correction next week. We will make every effort to do it as rapidly as possible, but I cannot offer any guaranteed delivery date. Along that line, a bug that is disabling will be attacked with immediate priority; a bug, that can be avoided by a change in a parameter card or by a "quick-and-dirty" fix of some variety, will receive a much lower priority. We'll tell everybody that the bug exists. We'll tell them how to get past it, and as soon as it rises to the top of the in-box, it will be fixed, and the permanent fix will be announced, but a non-disabling bug is going to be a relatively low-priority item.

Hickerson: Okay, are there any specific questions relative to Steve's comments? If there aren't, I would like to summarize and to discuss what (besides racing to the airport) we might do after leaving here. Archival automation is not an end in itself, and it has not transformed archival methodology nor changed our basic approach. It is affecting our methods of handling archival material, and there are problems, that we have ignored in the past, which become glaringly apparent when automation is applied. Frequently we have been able to pass over shortcomings in our manually-produced work which computer assistance makes more difficult to disregard. Our methods will also be altered by our now having the capability to provide certain types of detailed access that were not previously feasible. It was unreasonable to request that an individual arrange alphabetically 900 pages of index terms. We would have given them the terms and requested the results in a few years. A large computer can complete the alphabetization in a matter of seconds. Computer assistance may seem more expensive for some projects, but this is sometimes the result of our having never accurately measured the costs of manual production. It is quite clear, that for some archival tasks, a computer can produce a more accurate product, more quickly, and at considerable financial savings. In libraries the use of automation has reduced the cost of cataloguing only slightly, but the fact that a book may appear on the shelf one

or two years sooner is a substantial improvement. Like libraries, archives now have opened to them a new realm of possibilities which could not be accomplished previously.

I feel that this has been a useful meeting and that a community of users has been formed. The group is also large enough in number that it could carry out certain things in common. A few years ago there were but two or three successful applications, and people were saying,"Whatever happened to SPINDEX?" A users group would not have been viable at that point, but a lot of changes have taken place during the last two or three years. I would like to suggest that we now form a users group. Although I'm not sure of all that it might do, I hope that it might carry on, in some form, what has happened here during the last two days. Obviously, the NHPRC and the National Archives are going to do certain things. Concerning the design of a national guide, the NHPRC is in the prime position to perform that task. NARS is committed to providing responsible maintenance and development, but they are limited. Everyone is quite honest about the fact that while SPINDEX III is probably going to be quite beneficial to all of us, if the NHPRC guide project were not taking place, there wouldn't be a SPINDEX III. If a group of users needed a modification which didn't fit the needs of the National Archives, the modification would not be made by NARS. I am proposing that a users group be formed for three basic purposes. One is to organize communication: perhaps to print an occasional, inexpensive newsletter with technical notes. If you're doing an internal project description, you could also contribute it to the newsletter. You would not need to write the lengthy description that you would prepare for journal publication, but you could submit a brief project outline, identifying problems you have had or are having. We will find out who to call about problems that are new to us. We have made a big step in that direction already. (Once you get each other's phone numbers, it will probably ruin everyone's life, Steve's in particular.) Communication is very important. Second is the compatibility question we talked about yesterday: establishing a minimum common design for archival compatibility, or, as Steve said, compatibility between data bases which you may wish to integrate. When you don't expect to integrate, there is no need to worry about it. Third is future development. It well may be possible for this group to fund, on a contract basis, the creation of a new data entry program. It also might be possible to find external funding for common development. Those are the justifications I see at this point.

I would suggest, as a method of doing this, that we have a volunteer steering group which would, perhaps between now and the SAA meeting in Nashville [October 3-6, 1978], come up with a general definition of what this group is or might be. I am not sure that the group, if formed, would not disappear before too long. I have felt that, in the development of library automation, the formation of users groups performed some very specific developmental functions and had a strong effect on the larger profession. Particular groups played a role in spreading the "word" and in institutionalizing the "word" within the library profession. Likewise, we might play that same contributing role for the archival profession. So I suggest that we form a group of three or five steering committee members. They would specify a time at the next meeting of SAA at which we would have an organized meeting and perhaps a follow-up meeting. I'm sure we could meet at great length. So, I'm throwing

this suggestion out; we do not have a lot of time for duscussion. Are there comments concerning this? Yes.

Dennis A. Burton: I think that everything you mentioned is worthwhile, but, if nothing else comes out of it other than some form of communication, we can all save ourselves a lot of time and probably a lot of money and a lot of headaches just by knowing what other people are doing.

Charles H. Lesser: The minimum, I think, might be the annual gathering at SAA, which most of us go to anyway.

Frank G. Burke: Could I ask a question of Steve? Steve, what are the prospects, if one or more of the users developed a system modification that was acceptable to the National Archives as a useful addition to SPINDEX, that the new routine could be incorporated within the package itself and that NARS would provide maintenance on it?

Hannestad: I'm now getting into those grounds where I have been told by "mahogany row" that I must remain extremely cautious. (laughter)

Burke: We'll take a specific example and get back to the hoary old one that we always talk about, the alpha-sort package.

Hannestad: Okay. If a users group, either a formal or informal users group, decided, collectively, to contract for program development in a specific area, and it was a developmental area in which the National Archives saw some potential use, then, I believe that the National Archives would be willing to pay its proportionate price for that routine. In other words, the same way that we are selling to you, the users, SPINDEX III's enhancements over SPINDEX II at some cost, if some other group than the National Archives develops programming that, we find, could be useful to us, then we would, on an individual-case basis, determine whether we wish to purchase that option too. We would not necessarily state that we demand that you give it to us for free; on the other hand, if you did offer it to us for free, we would gratefully accept it. (laughter) The concept of a users group has, Frank, you can speak to this, come up at each of the SPINDEX users conferences, and I think that each time the concensus opinion has been that it's a good idea, and each time we have had to state that, though the National Archives looks upon the formation of a user group as a desirable goal, we, unfortunately, are not in a position that we could provide the secretariat for it or provide funding for it. I'm speaking for the National Archives itself, not necessarily for the Commission [National Historical Publications and Records Commission]. I would have no idea whatsoever how the Commission might react to a proposal made to it, but the National Archives, as the National Archives, is not in a position, today,

where it could support a users group, either by providing a secretariat or a development staff.

Brent G. Thompson: What about the possibility of having a small amount of dues for covering things like newsletter costs, two, three dollars a year? If someone would be willing to provide free secretarial help, that would pay for the printing and mailing of a small newsletter.

Burke: Since we're all with organizations and institutions that find it sometimes difficult to pay dues to something like that, it might be better if the newsletter had a charge to it so that the institutions actually bought it. The library would then buy this thing, and the money would go back instead of being on a dues paying basis. I think that there would be other people, who are not in this group, who would probably want to buy it also, just because they buy everything in the field. (laughter) Well, that's true. There are people who, if it says "information" or says "genealogy", they buy it. (I'm not mixing the two.)

Hickerson: Yes.

Richard S. Parmenter: Steve, can you tell us how many copies you have sold, and how many of those people are represented here?

Hannestad: I really wish I could answer that question. Unfortunately, I can't, because, for a period of approximately two years, we lost control over our sales, and copies of the tape went to places that we do not know about. Incidently, this is one reason why we refused to participate in the federal software exchange, which allows copies to be sold without our ever knowing about it. Right now I know of one recent purchaser, Northern Illinois University, that is not represented here. Hoover Institution [Hoover Institution on War, Revolution and Peace] has, at various stages of the game, been involved, and they definitely have a copy of some version of the system. I have no idea which version they have. We believe that Wisconsin [State Historical Society of Wisconsin] has a version of the system; it is possibly a DOS version. They don't know what version they have, and we don't know what version they have.

Burke: Minnesota [Minnesota Historical Society] has a version.

Hannestad: So we're in a situation where we are not positive who all is involved. We have industrial customers, and I am always a little bit hesitant to say what industrial customers have purchased it, if they do not choose to participate in the users associations. Frankly, some of them have just been "black

box" purchases. We have sold them the system; they have not told us what they intend to us it for, or anything else, and have, very specifically, said that they were not interested in participating with others. So there are probably at least six more users than are represented today, and there are people that have tapes but cannot really be called users.

Burke: I think that, from our feel for the situation, if there was a move toward some coordinated effort on the part of the present users, and there was a move toward the development of a common use of the program, that there are, in the wings, a number of people and institutions that are ready to jump in that direction.

Hickerson: I think that the beginning user or potential user may receive as much benefit as anyone, as a result of our providing them with a point of access to information. The National Archives is not going to be able to fill all information needs. People need someone to contact concerning the archival aspects of their application. Concerning the question of dues, although it's probably better to sell our publication, I think that the idea of people actually making a financial commitment is of some significance. I don't think that, at this point, we should start discussing a dues structure, but it is a way of getting a little leverage, and I would suggest that, if people don't expect to get a few dollars worth of benefit from belonging, then, they probably won't. Ron.

Ronald G. Watt: You made a suggestion about a steering committee, and we've sort of deliberated around that. Coming back to that, I'm just trying to think of institutions which could be represented on that steering committee and could give us a report at the next SAA meeting. I would suggest three or four institutions: one would be NARS; one would be NHPRC; I think another should be yourself [Cornell University]; and another should be Dick Parmenter from INCO. That's four. I think those four institutions could select their representative for the steering committee, and it could certainly give us a good report at SAA.

Burke: Steve, would you feel that was redundant or would you feel for any reason that--

Hannestad: You mean, to have both the Commission and NARS represented?

Burke: I would feel very comfortable serving on the steering committee; I don't know how comfortable Steve would be.

Hannestad: I would be reluctant to speak for the Commission, but I would not be reluctant to have the Commission speak for NARS.

Watt: I really think it would be valuable though to have Steve on that steering committee.

Hannestad: I'm not going to say, "If nominated I will not run; if elected I will not serve." (laughter)

Hickerson: Well, I would like to add someone from the West, and, obviously, the people at LDS have the most experience at this particular point in time.

Burke: I would agree.

Hickerson: Is there anybody else who would like to serve?

Burke: This steering committee might, I presume, be disbanded after October.

Hickerson: I suggest that the primary task of this steering committee should be to provide the format for the meeting in October. It would be important for somebody to be responsible for coordinating the work of the committee, and I don't want to do it. Once again, I suggest LDS. What do you think of that?

Watt: I don't know. (laughter)

Burton: While someone's thinking on that, I think you should consider South Carolina too.

Hickerson: That's quite appropriate, and, especially, if we only have one representative from the NARS structure, which, in some ways, I think is desirable.

Hannestad: The point of the matter is: our offices are only several blocks apart, and we have been known to drink coffee with each other, so, a single representative of NARS would be appropriate. Although we are not, officially, both represented, unofficially, you can be certain that we would be working with each other, and I think that South Carolina might be a very legitimate addition to it. One question I feel somewhat compelled to bring up: most of the people here are indeed archivists or records managers; I'm just won-

dering though, if putting this as an adjunct to the SAA might not be working a hardship on some of the users. I don't know whether we could be excluding a portion of our users, STELCO, INCO, Coca-Cola, who are doing some of the most innovative work with SPINDEX.

Hickerson: I see that as a very real problem. On the other hand, suitable alternatives are not very easy to come up with.

Burke: Coca-Cola invariably shows up at SAA meetings.

Louise Lloyd: SAA includes records managers and certainly we need to emphasize records management in it.

Hickerson: I think that in a lot of state archives, if they adopt SPINDEX, they will want to use it as a records management system. Apparently the commercial applications are going to continue to expand. Strangely enough, it's getting to be a proven product.

Arthur D. Dick: It seems to me that the concept of a newsletter is going to be more than adequate for our purposes. If we feel we need to participate, we'll be there. We too have an Office of Management and Budget that controls what we do and when we do it. So, I think you should go ahead, and I think that the archivists, the SAA people, are still going to be your primary group. From my point of view, I won't hesitate to get on the phone as soon as I get your telephone directory, if there's some need for us to do that, and I'm prepared to submit to a newsletter those things that we're doing, that we think are creative.

Marvin Y. Whiting: Tom.

Hickerson: Yes.

Whiting: It seems to me that one thing that has not been mentioned is, if you do have a steering committee preparing for some sort of meeting at SAA, that there will be expenses entailed in mailings, which will have to be borne somewhere.

Burke: Well, it isn't that much of a burden with this size group.

Lesser: Fifteen letters, twenty letters, that's not so bad.

Burke: If we're represented institutionally, I think we can get away with that. Since we're among friends, I feel that I can speak openly. I think that it's important not to have two NARS people on the committee, because it looks like a power play. Secondly, I am quite willing to serve on the steering committee. I think, again for outside appearance, that I'd rather not serve as the chairman or coordinator or whatever, but I'd be happy to serve and lend whatever help we can from our office, because it is important to us from a purely personal point of view, I mean from the view of the Commission, to have something where we can have some coordination.

Hickerson: Well, do people feel comfortable, as comfortable as they're going to get in the next couple of minutes, (laughter) with LDS, Cornell, MHPRC representing NARS generally, South Carolina, and--

Burke: INCO.

Hickerson: --and INCO. Let me, real quickly, readdress my earlier question to Salt Lake, Would you be interested in playing the coordinating role, at least until October?

Watt: I guess we could do it.

Hickerson: Okay, I admit you were put on the spot, and you will always be able to use that as a disclaimer. (laughter)

Cleo A. Hughes: Do you want to run it as an ad hoc committee [of SAA] and see if we can run it through the Chicago office as an ad hoc committee--

Burke: I'd rather not.

Hughes: --so that all of the information could be printed in the SAA Newsletter? They could provide extra copies to us, that we could mail to the people who are not SAA members, and put it on the program for Tuesday.

Hickerson: I don't know at this particular point.

Hughes: You think it would attract too many people who don't have SPINDEX, or what?

Hickerson: Well, at this particular point, SAA is also dealing with what may become an ad hoc committee on national bibliographic system needs. Perhaps this might confuse that issue, which I think is a separate issue. Although I believe that this group should be involved in those questions, I would not want to see those two matters confused, or it appearing that one is dominating the other.

Burke: Also, on an ad hoc committee, the Council has the right to select the chairman, and what's the rule Dick, about Council members serving as committee members?

Richard H. Lytle: Committee membership is fine, but you're not supposed to serve as a committee chairman.

Hickerson: I believe that we had better get on our way as quickly as possible. I thank you; I'll send you little bits of the trailing fragments of the conference shortly. Thank you.

STATE OF WASHINGTON
HISTORICAL RECORDS SURVEY

☐ PUBLIC ☐ PRIVATE
☐ COL/RG ☐ SUB-GROUP ☐ SER

INSTITUTION/ORGANIZATION & CITY	DIVISION/OFFICE OF RECORD	SURVEYOR

OFFICE USE ONLY	ST	CIT	INS	COLL	SER	LEVEL	

TITLE

VARIANT TITLE(S)

CONTENTS

☐ AGREEMENTS
☐ BIDS
☐ BUDGETS
☐ CAMPAIGN MATERIALS
☐ CERTIFICATES, STOCKS & BONDS
☐ CHARTS
☐ CONTRACTS
☐ CORRESPONDENCE
☐ COURT DOCUMENTS
☐ DEEDS
☐ DIARIES
☐ DIRECTIVES
☐ DRAWINGS
☐ EPHEMERA/PRINTED MATERIALS/PAMPHLETS
☐ FINANCIAL RECORDS
☐ JOURNALS
☐ LEDGERS
☐ LEGAL DOCUMENTS
☐ LEGISLATIVE MATERIALS
☐ MAPS
☐ MEMORANDA
☐ MINUTES
☐ NEWSPAPER / CLIPPINGS
☐ ORDINANCES
☐ PHOTOS
☐ PLANS
☐ PLATS
☐ REPORTS
☐ RESOLUTIONS
☐ SCRAP BOOKS
☐ SEE TITLE
☐ SPEECHES & WRITINGS
☐ UNPUBLISHED MANUSCRIPTS
☐ VOTER RECORDS
☐ CHURCH RECORDS (SPECIFY) _____
☐ OTHER (SPECIFY) _____

QUANTITY

VOLUMES	CUBIC FEET	LINEAR FEET	UNITS OF MICROFORM
V	C	L	M

SPAN DATES				BULK DATES			
FROM	YR. / MO.	TO	YR. / MO.	FROM	YR. / MO.	TO	YR. / MO.
	/		/		/		/

DESCRIPTION/FUNCTION/ARRANGEMENT/PRIMARY SUBJECT CHARACTERISTICS

CHARACTERISTICS	STATUS	ARCHIVAL CUSTODY (INACTIVE, DEPOSITED MATERIALS)
☐ ORIGINALS ☐ REPRODUCTIONS ☐ BOTH	☐ ACTIVE ☐ INACTIVE	☐ YES ☐ NO

RESTRICTIONS (SPECIFY, IF ANY)

FINDING AIDS
☐ CARD CATALOG ☐ CONTAINER LIST ☐ INDEX
☐ INVENTORY ☐ REGISTER ☐ NONE ☐ OTHER (SPECIFY) _____

DATE INFORMATION COLLECTED	TYPE OF INSTITUTION/POLITICAL UNIT
YR / MO / DY	☐ STATE ☐ COUNTY ☐ CITY ☐ CHURCH ☐ LIBRARY ☐ BUSINESS ☐ PRIVATE PERSON ☐ HISTORICAL SOCIETY/ MUSEUM ☐ OTHER (SPECIFY) _____ ☐ SPECIAL DISTRICT (SPECIFY) _____

COUNTY OF LOCATION	COUNTY/STATE/ NATION OF ORIGIN	MEDIUM
		☐ ADP ☐ AUDIO TAPE/PHONO ☐ CARTOGRAPHIC ☐ VIDEO TAPE/MOVIE ☐ PAPER ☐ PHOTOGRAPHIC ☐ MICROFORM (SPECIFY) _____

BIOGRAPHICAL/HISTORICAL INFORMATION/COMMENTS	OFFICE USE ONLY

112

WSHRAB FORM 4

*Reduced from 8-1/2" x 14"

Appendix B

NHPRC Guide Work Sheet

TAG	LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 8
000	Control	Control	Control	Control
10x	Index references	Index references	Index references	Index references
110	Parent institution	Title	Title	--
120	Institution unit	Type material	Type material	--
130	Street address			
140	Mailing address			
15				
160	Telephone no.			
17				
180	City	--	--	--
190	State	--	--	--
200	Service hours			
210	User fees			
220	Access policy			
230	Copying facilities			
24x	Acquisitions polity			
25				
26				
27				
28				
29				
300	Quantity	Quantity	Quantity	
310	Beginning date	Beginning date	Beginning date	
320	Ending date (span d.)	Ending date	Ending date	
33x	Description	Description	Description	
340		Restrictions	Restrictions	
350		How acquired	How acquired	
360		Accession no.	Accession no.	
370		NUCMC no.	NUCMC no.	
380	Guide title	Finding aids	Finding aids	
39x	Bibliography			

TAG	LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 8
400	Project code	Project code	Project code	Project code
41x	Index, Repository	--	--	Level 1 overflow
42x	Index, Organization	--	--	Level 1 overflow
43x	Index, Geographic	--	--	Level 1 overflow
44x	Index, Subject	--	--	Level 1 overflow
45x	Index, Acquisitions	--	--	Level 1 overflow
46x	Index, Personal name	--	--	Level 1 overflow
47	--	Importance code	Importance code	
48	--	Spin-off code	Spin-off code	
49	Repository type			
90x	Abstract	Abstract	Abstract	
91				
92				
93				
94				
95				
96				
97				
98				
99				

DEPT/SECT RECORDS CENTER HOLDINGS SERIES

02/14

---COMPTROLLER'S/DATA CONTROL

* Z0169 Accounts Payable Cheque Reconciliation Cards
OFF:1mo RC:2mo TOTAL:3mo

YEARS	RANGE	BOX #	REF.	REC'D	DESTROY	
						Z0169-A
						Z0169-B
						Z0169-D
						Z0169-E
75-75	Sept.25-Oct.7	013-25		10SEP75	1976/02	Z0169-001
75-75	November-12	013-31		10SEP75	1976/02	Z0169-002
75-75	December	008-30		18DEC75	1976/04	Z0169-003
75-75	November	008-34		22DEC75	1976/04	Z0169-004

* Z0177 Contract Receipts (Tab Cards)
OFF:1mo RC:3 TOTAL:C+3

YEARS	RANGE	BOX #	REF.	REC'D	DESTROY	
						Z0177-A
						Z0177-B
						Z0177-D
						Z0177-E
74-74	January	001-01		14NOV75	1978/01	Z0177-001
74-74	January	001-03		14NOV75	1978/01	Z0177-002
74-74	February	001-02		14NOV75	1978/01	Z0177-003
74-74	February	001-04		14NOV75	1978/01	Z0177-004
74-74	March-April	001-05		14NOV75	1978/01	Z0177-005
74-74	April-May	001-06		14NOV75	1978/01	Z0177-006
74-74	May-June	001-07		14NOV75	1978/01	Z0177-007
74-74	July-August	001-08		14NOV75	1978/01	Z0177-008
74-74	August-September	001-09		14NOV75	1978/01	Z0177-009
74-74	October-November	001-10		14NOV75	1978/01	Z0177-010
74-74	November-December	001-11		14NOV75	1978/01	Z0177-011
73-73	January	001-13		14NOV75	1977/01	Z0177-012
73-73	January-February	001-15		14NOV75	1977/01	Z0177-013
73-73	February	001-14		14NOV75	1977/01	Z0177-014
73-73	March	001-16		14NOV75	1977/01	Z0177-015
73-73	March-April	001-17		14NOV75	1977/01	Z0177-016
73-73	May	001-18		14NOV75	1977/01	Z0177-017
73-73	June	001-19		14NOV75	1977/01	Z0177-018
73-73	July-August	001-20		14NOV75	1977/01	Z0177-019
73-73	August-September	001-21		14NOV75	1977/01	Z0177-020
73-73	October	001-22		14NOV75		Z0177-021
73-73	November	001-23		14NOV75	1977/01	Z0177-022
73-73	December	001-24		14NOV75	1977/01	Z0177-023
72-72	July	001-25		14NOV75	1976/01	Z0177-024
72-72	August	001-27		14NOV75	1976/01	Z0177-025
72-72	September	001-26		14NOV75	1976/01	Z0177-026
72-72	October	001-28		14NOV75	1976/01	Z0177-027
72-72	October-November	001-29		14NOV75	1976/01	Z0177-028
72-72	December	001-30		14NOV75	1976/01	Z0177-029
75-75	January	016-13		14NOV75	1979/01	Z0177-030
75-75	February	016-15		14NOV75	1979/01	Z0177-031
75-75	March	016-14		14NOV75	1979/01	Z0177-032
75-75	April	016-16		14NOV75	1979/01	Z0177-033

YEARS	RANGE	BOX #	REC'D	RS #	DEPT
1976/08					
75-75	June 15 - July 27	059-30	01APR76	Z0050	02/18
73-73	June-July	032-42	28NOV75	Z0150	02/10
71-71	July 1-6	021-05	05DEC75	Z0174	02/10
71-71	July 7-12	021-07	05DEC75	Z0174	02/10
71-71	July 13-18	021-09	05DEC75	Z0174	02/10
71-71	July 19-24	021-11	05DEC75	Z0174	02/10
71-71	July 1-2	021-01	05DEC75	Z0174	02/10
75-75	July	003-21	17NOV75	Z0179	02/14
1976/09					
75-75	August 3-31	059-31	01APR76	Z0050	02/18
73-73	May-June (Monthly)	036-39	28NOV75	Z0152	02/10
73-73	July-August (Monthly)	036-45	28NOV75	Z0152	02/10
71-71	July-Aug 02-99	022-46	10DEC75	Z0171	02/10
71-71	August 1-6	021-06	05DEC75	Z0174	02/10
71-71	August 7-12	021-08	05DEC75	Z0174	02/10
71-71	August 13-16	021-10	05DEC75	Z0174	02/10
71-71	August 19-22	021-12	05DEC75	Z0174	02/10
71-71	August 1-2	021-02	05DEC75	Z0174	02/10
75-75	August	003-19	17NOV75	Z0179	02/14
1976/10					
73-73	Jan-Sept (By Qtr.)	032-38	28NOV75	Z0150	02/10
73-73	August-Setember	032-48	28NOV75	Z0150	02/10
73-73	Jan-Sept (By Qtr.)	037-43	28NOV75	Z0152	02/10
73-73	July-September	037-42	28NOV75	Z0154	02/10
73-73	July-September	032-43	28NOV75	Z0150	02/10
71-71	September 02-99	022-48	10DEC75	Z0171	02/10
75-75	September	003-23	17NOV75	Z0179	02/14
75-75	July-Sept.	005-20	17NOV75	Z0183	02/14
1976-11					
75-75	January-April	032-27	22JAN76	Z0041	02/18
75-75	Sept. 7 - Oct. 5	059-32	01APR76	Z0050	02/18
73-73	September-October	036-41	28NOV75	Z0152	02/10
71-71	October 1-2	033-01	05DEC75	Z0174	02/10
71-71	October 1-6	033-05	05DEC75	Z0174	02/10
71-71	October 7-12	033-07	05DEC75	Z0174	02/10
71-71	October 13-18	033-09	05DEC75	Z0174	02/10
71-71	October 19-24	033-11	05DEC75	Z0174	02/10
75-75	October	003-24	17NOV75	Z0179	02/14
75-75	October	005-16	17NOV75	Z0183	02/14
75-75	October	008-35	19DEC75	Z0186	02/14
1976/12					
75-75	Oct. 12 - Nov. 9	059-33	01APR76	Z0050	02/18
73-73	November (Monthly)	036-47	28NOV75	Z0152	02/10
71-71	Oct-Nov 02-99	022-44	10DEC75	Z0171	02/10
71-71	November 1-6	023-06	05DEC75	Z0174	02/10
71-71	November 7-12	023-08	05DEC75	Z0174	02/10
71-71	November 13-18	023-10	05DEC75	Z0174	02/10

TITLES	RS #	LOC	D	C	RC	OFF	RC	T	START-TERM	REMARKS
ACCT N-AC CONTINUED										
Purchase Journal Entries	N1314	NYO		X	X	2	5	7		Review with Tax Section Before Destruction
Purchase Orders	N1388	NYO		X	X					
Reference Individual Reference Files	N1300	NYO			X					Sub-Schedule
Reference Section Reference Files	N2031	NYO								Sub-Schedule
Renegotiation Act	N1346	NYO		X	X	1	6	7		
Retirement System - Cancelled Checks	N1775	NYO		X	X					
Retirement System - Mortgage Files	N2793	NYO		X	X					
Retirement System Files	N2808	NYO		X	X					
Retirement System Ledger	N2775	NYO		X	X					
Retirement System Normal Provision	N1389	NYO								
Royalty Files	N1295	NYO		X	X					
Sales	N1434	NYO		X		5		5		
Sales Registers - Miscellaneous Customers	N1679	NYO		X	X					
Section Reference Material	N2025	NYO		X		P				Manuals, Duplicate Set Year End Inventories Plant Chart of Accts GGH - Govt Contracts
Securities Statements	N1787	NYO		X	X					
Security Advices - Corporate	N1390	NYO		X	X					
Sickness & Disability Records with Bank Reconciliation & Cancelled Checks & Vouchers	N1536	NYO		X	X					
Sickness Disability Cash Book	N1950	NYO		X	X					
Source Data Customer Sales Orders-Completed	N1656	NYO		X	X	6mo	2.5	3		
Source Data Daily Advice of Shipments	N1683	NYO		X		1		1		
Source Data Drawback Prices	N2028	NYO		X		4		X		Review by Senior Sales Acct Before Destruction
Source Data Invoices	N1298	NYO		X	X	1	5	6		Inco Inc & Inco Canada Excluding HAPD Royalties
Source Data Summary of Shipments and Receipts	N2019	NYO		X		C+1		C+1		Inco Canada
Special Costs & Statistics	N2023	NYO		X	X	P				Sub-Schedule
Special Order Expenditure Statements	N1677	NYO		X	X					
Special Order Expenditures	N1485	NYO		X		C+2		C+2		
Special Order Requisitions	N1486	NYO		X	X	T	15	T+15		
Specimen File	N1424	NYO		X		5		5		Blank Forms
Statutory Account for Mond and Subsidiaries	N1391	NYO								

RETENTION POLICY INDEX WORKING COPY

TITLES	RS #	LOC	D	C	RC	OFF	RC	T	START-TERM	REMARKS
ADMINISTRATIVE										
Administrative File	N1156	NYO		C		X		X	1966-	
Individual Reference Files - Administrative Services	N1191	NYO	D			X		X	1965-	
Administrative Files	N1210	NYO								
Office Administrative Files	N1431	NYO		C		T		T		
General - Administrative (Inco) - Washington District Office	N2949	NYO		C						
General - Administrative - Washington District Office	N2950	NYO		C						
Capital Appropriation Request - Administrative Services	N2980	NYO		C		C+5	5	C+10	1972-	
Short Term Promissory Notes - Inco Canada Issue - Administrative and Policy Correspondence	T3029	TOR		C		5	*5	*10	-	Retention Approved 01SEP7[illegible]
Bank Lines of Credit - Administrative and Policy Correspondence	T3066	TOR		C		5	*5	10		Retention Approved 01SEP7[illegible]
ADMS										
Project and Programs - ADMS. Master File	N2801	NYC	D			P				
ADVERT.SEMENTS										
Periodicals for Advertisements - Past or Potential	L0553	LON		C	RC	2		*2		
Periodicals with Inco Advertisements	L0554	LON		C		C		C	-1969	
Advertisements - Inco Inc. - Bound Books	T4306	TOR			RC	Awaiting Approval			1959-1962	
Advertisements - Canadian - Bound Books	T4308	TOR			RC	Awaiting Approval			1933-1964	
Advertising										
Advertising Pulls	L0541	LON		C		*2				
Correspondence - Horizontal Advertising	L0543	LON				1	1	2	-1971	
Current Advertising Yearbooks	L0547	LON		C	RC	1	1	2		
Technical General Advertising Schedules	L0556	LON		C	RC	1	2	3		
Advertising Summaries - Inco Inc.	N1822	NYO		C	RC					
Advertising	T4080	TOR		C	RC	1	5	6	1965-	Retention Approved 01AUG74
Reports - Advertising	T4307	TOR			RC	Awaiting Approval			1965-1970	
Advice										
Stock Advice Notes	L0199	LON			RC		5	5	-1967	
Advice Notes Clydach	L0300	LON			RC		6	6	-1967	
Source Data Daily Advice of Shipments	N1683	NYO		C		1		1		
Technical Advice	T4003	TOR		C	RC	1	9	10	1965-1973	Retention Approved 01JUN72
Advices										
Cash Receipt Advices; Bank Shipping Letters & Acknowledgements	L0301	LON		C	RC	1	1	2		
Delivery Advices	L0303	LON			RC		5	6	-1967	
Security Advices - Corporate	N1390	NYO		C	RC					
Daily Cash Receipts - Customer Advices	N1678	NYO		C	RC					
Advisory										
Ret Sys - Advisory Fee File	N1355	NYO		C		P				
AFL-CIO										

DOC NO.	BIBLIOGRAPHIC ENTRY	LOCATION
AIR	CONTINUED	
71-428	The Value of Time in Passenger Transportation: The Demand for Air Travel., R. Gronau. 1970 National Bureau of Economic Research, 74 P.	NYO N-BL 420 G
71-356	Air Quality Act of 1967, Hearings Before Committee of Interstate and Foreign Commerce, House of Representatives, 90th Congress, 1st Session on H.R. 9509 A Bill to Amend the Clean Air Act..., S.780 An Act to Amend the Celan Air Act..., August 15, 16, 17, 18, 21, 22, 23, 24, 1967, Govt. Printing Office, 887 P.	NYO N-LL 822.001 H
71-378	Air Quality Act of 1967 Amending the Clean Air Act as Amended, Report of the Committee on Public Works. U.S. Senate to Accompany S. 780, 90th Congress, 1st Session, 1967, Govt. Printing Office, 86 P.	NYO N-LL 822. 001 S
AIR QUALITY		
71-356	Air Quality Act of 1967, Hearings Before Committee on Interstate and Foreign Commerce, House of Representatives, 90th Congress, 1st Session on H.R. 9509 A Bill To Amend the Clean Air Act..., S.780 An Act to Amend the Clean Air Act..., August 15, 16, 17, 18, 21, 22, 23, 24, 1967, Govt. Printing Office, 887 P.	NYO N-LL 822. 001 H
AIR QUALITY ACT		
71-378	Air Quality Act of 1967 Amending the Clean Air Act as Amended, Report of the Committee on Public Works. U.S. Senate to Accompany S.780 90th Congress, 1st Session, 1967, Govt. Printing Office, 86 P.	NYO N-LL 822. 001 S
AIR QUALITY ACT OF 1967		
71-356	Air Quality Act of 1967, Hearings Before Committee on Interstate and Foreign Commerce, House of Representatives, 90th Congress, 1st Session on H.R. 9509 A Bill to Amend the Clean Air Act..., S.780 An Act to Amend the Clean Air Act..., August 15, 16 17, 18, 21, 22, 23, 24 1967, Govt. Printing Office, 887 P.	NYO N-LL 822. 001 H
AIRLIE		
70-564	Libraries and Automation. Proceedings of the Conference on Libraries and Automation	NYO N-BL 350 M

PRES. NO. SLIDE PRESENTATIONS KEYWORD INDEX

EXPANSION
N75- 8.05 Shimura Kaku - Tokyo Nickel Expansion Plans, C.E. O'Neill 18/AUG/75
N75- 8.05 7. Plant Expansion Nickel Oxide Sinter 75, Sulfuric Acid, Nickel Pig
N75- 8.05 8. Plant Expansion Flowsheet
N75- 8.05 9. Estimated Capital Costs, Plant Expansion
N75- 8.05 10. Estimated Operating Costs, Plant Expansion

EXPENDITURE
N75- 3.06 4. Base Case Capital Expenditure Assumptions
N75- 5.10 9. Analysis of 1975 Capital Expenditure Savings

EXPENDITURES
N75- 4.06 1. Exploration Department - Total Exploration Expenditures
N75- 4.06 2. Exploration Department - Mines Exploration Expenditures
N75- 4.06 4. Exploration Department - Field Exploration Expenditures
N75- 5.10 8. Capital Expenditures
N75- 6.04 27. Exploration - Forecast Expenditures
N75- 6.04 28. Exploration - Forecast Expenditures - Field Exploration by Commodities
N75- 6.04 36. Mines Exploration - Ontario Division Forecast Expenditures
N75- 9.02 14. Capital Expenditures
N75- 9.12 14. Capital Expenditures

EXPENSE
N75- 3.05 2. Ontario Division Operating Expense
N75- 5.06 2. SG&A Expense
N75- 5.06 3. Exploration - Research and Development - Interest Expense
N75- 5.10 6. Expense Analysis - Metals Business - By Class
N75- 5.10 7. Expense Analysis - Metals Business - By Company

EXPENSES
N75- 5.06 4. Other Expenses - Freight and Insurance - Other Income
N75- 5.10 4. Primary Metals and Rolling Mills Cash Operating Expenses - By Class

EXPLANATORY
N75-10.05 Explanatory Financial Section

EXPLORATION
N75- 4.06 ESB - Inco Roundtable, W. Steven - Mineral Exploration and New Developments, Skytop, 01 - 02/APR/75
N75- 4.06 1. Exploration Department - Total Exploration Expenditures

Appendix D

George E. Mariz, "SPINDEX II at Western Washington University" [This paper was contributed to Session II but not presented verbally.]

In May, 1976, the National Historical Publications and Records Commission awarded the Western Washington University Graduate Program in Archival and Records Management Administration a grant to survey the records of Public Utility Districts in Washington and Oregon. From the survey were to emerge a listing of the historically important resources of the agencies as well as various documents to aid the PUDs in records management. The proposal's authors committed the project to the use of the SPINDEX II system in the preparation of all the survey's products.

The variety of records surveyed and the somewhat unusual demands made of the system offered problems and opportunities beyond those which collections of archival material generally present. The materials were diverse in terms of function, physical characteristics and location. In addition, we undertook a project concerned as much with records management as with archives. At the same time there were fewer problems in other ways. The work would require very little in the way of sophisticated indexing capability as there were very few records identified by means other than a record series title. Names of individuals did occur in some cases, but the survey dealt largely with a relatively limited number of terms, certainly many fewer than the Cornell project encountered. The keyword indexer program generated slightly more than two thousand one hundred items with the most frequently occuring word, 'files,' appearing over six hundred and fifty times. Over six hundred terms appeared only once.

The form, which is appended, is relatively simple. It uses SPINDEX's built-in flexibility, creating several fields for use in compiling data on record volume, physical characteristics, legal requirements, possible scholarly value and series inconsistencies.

Forms were completed in the field by previously trained surveyors and then mailed to the project office. There project staff checked the arrangement of forms and coded them. They were then passed on to other project staff who inspected them for errors by field surveyors or office personnel. This worked well when project staff-time was available, but the rate of error and attendant problems increased as the skill of available help declined.

The forms were then given to data processors who entered the information on cards. After project staff cursorily inspected the cards, largely to check for easily recognizable errors, the data was entered on master files.

Beginning in June, 1977, project staff began to produce registers, including an index, from the data. Survey personnel continued to process and enter data until December, 1977, and then produced a full print out and index of the material contained on approximately forty thousand forms. (As an interesting side-light the survey has produced over twenty cubic feet of records during its course: ten cubic feet of survey forms, more than seven cubic feet of punched cards and approximately four cubic feet of computer output.) The four registers and index run to approximately two thousand four hundred pages.

Our data base included over one-half million items.

The survey determined the tag structure, constructed the form and made initial runs of data in October, 1976, using material gathered from the files of the Bonneville Power Administration relating to the Washington and Oregon PUDs.

Control Field-Tag 000

This field serves as a record identification number. It consists of a four digit record group number in blocks 1-4 which is the identifying number for the agency. Each public power agency, PUD, cooperative utility and municipal system in the three-state survey area has been designated with a code number. Blocks 5 through 8 contain a four digit office code. This is determined by referring to organizational charts used by the individual PUDs or which the interns created on-site while conductirg the survey. An agency history form was created for each PUD and was assigned an office code of 0000. Major departments received code numbers in the hundreds: Boards of Commissioners were 0100, Managers were 0200, Treasurers/Comptrollers were 0300 and so on. Subordinate offices were assigned codes in the tens within the range of its major department. Thus the auditor, a subordinate of the manager, was assigned office code 0220.

Blocks 9 through 12 contain a four digit record series code increasing by ten with each successive series: the first record series is 0010, the second 0020 and so on. Agency history and office history forms were coded 0000 in this space. If we discovered during processing that a series was misplaced or assigned to the wrong office, it could be placed in proper sequence by assigning it a series number which used the ones column: a series occuring between 0030 and 0040, for instance, could be assigned number 0031 or any other number between the two series.

Blocks 13 through 18 were used as a dummy.

Block 19, the record action code, was a one digit code used to control indentation of the agency title, office title or series title on the printed register. Agency history forms were coded with a 1, office history forms with a 2 and series description forms with a 3.

Block 20, the field action code, was used only for additions and corrections and during normal processing was left blank. When used during the correction procedure it was filled in with an A to add a record or a C to substitute the contents of a new record for an old one.

Series Title-Tag 110

As our record series titles were of a specific nature, and because we desired permutation of all words (except articles) in this tag field for indexing purposes, we used Tag 110 rather than Tag 100. On agency and office history forms, all words in the title were linked with an underscore as we did not desire their permutation. Series titles were limited to fifty-seven characters. This proved adequate for all but seven record series titles, and in no case did it cause confusion in the identification of records.

Record Characteristics-Tags 200, 201, 202 and 203

These fields were used to identify the contents, arrangement, size and

physical characteristics of the surveyed records. These blocks were not completed on agency or office history forms. In all fields more than one box could be checked, and in Tag 203, at least one box had to be checked in each column. All the material indicated was then punched on data cards. Entries appeared in a consecutive, vertical listing on the printed register. One egregious error we discovered was our failure to include a box for newspaper clippings. Over one thousand entries of this sort appeared in the 'other' category of Tag 200, far exceeding all other entries under this heading.

Inclusive Dates-Tag 250

This tag was used for all forms whether agency or office history or record series title. In the 'from' portion the earliest date of the record was coded using the four digits of the year to the left of the slash mark and two digits for the month if necessary. For example, February, 1936 became 1936/02; 1938 became 1938/00. Current series were indicated by 9999/00 in the 'to' block.

Record Quantity-Tag 300

We recorded the quantity of material in an agency, office or series with only one of the three categories indicated: cubic footage, linear footage or rolls of microfilm. For agency and office history forms quantities were expressed in cubic feet only. Elaborations of the general contents were made in Tags 200, 201, 203, 204 and the descriptive field in Tag 900. Cubic feet and linear feet were entered to the nearest tenth of a foot. Interns were instructed also to fill in decimal points and provide a space for them. Quantities appeared as follows: 35.5 cubic feet became C035.5; 2 linear feet became L002.0; 5 reels of microfilm became M00005.

Location Code-Tag 400

This was a ten digit code used to give the location of the agency, office or series. Each site for each PUD was assigned a separate number. The location code consisted of the four digit RGN, which is the same as the four digit code in Blocks 1-4. Thus the agency code number and first four digits of the location code are always the same. This was followed by three zeros and a three digit serial number beginning with 001. Each separate physical location holding records within a single agency received a code number.

Record Classification-Tag 500

This tag was created specifically for use in the state of Washington, the only one in the area which uses OPR (Official Public Record) and OFM (Office Files and Memoranda) designations. Official Public Records document financial transactions of public agencies, bonds, claims by the agency against state and local government or private parties or claims against the PUD. It also includes any records the PUD by law is required to keep or file with supervisory or regulatory agencies. Office Files and Memoranda are those records which do not fall under OPR classification and are used in the internal administration of an agency. We have decided to discontinue use of this tag field for two reasons. First, it pertains only to one state in the survey. Second, even in Washington it was frequently difficult to determine if a record was in one or the other classification.

Determination of Archival Value-Tag 550

Again, this was an added field which arose from the need to have interns who were trained historians and apprentice archivists judge the ultimate value of a series to the government, the originating agency and scholars. Interns were advised to consider the possible value of a series in light of their experience with historical sources, use to the agency and legal requirements. Generally interns (true archivists all!) erred on the side of caution and declared archival value for many series which certainly will be disposed of when time for appraisal comes.

Recommended Retention-Tags 600, 601 and 602

Yet another added set of fields, these were incorporated into the form to allow us to make decisions on the scheduling of records. In tag field 600, interns were instructed to enter the retention period specified by statute or a records schedule if it existed. In tag field 601 the intern entered the retention period the agency used. The surveyor was instructed to enter his opinion for a retention period in tag field 602. Obviously, a surveyor's decision on the 550 tag field had an important influence here. In many cases surveyors and organizational officers did not know the legal period of retention and many PUDs had never made decisions on records disposition.

Series Inconsistencies-Tag 700

This field was created to allow us to discover if small or significant portions of a series were missing, or if record keeping had been altered for any reason. Surveyors were instructed to check the File Break' box if the series were not continuous between the inclusive dates shown in Block 250 (i.e., a missing year or month). They were to check 'Missing Items' if individual folders had been removed which did not offset the overall continuity of the series. Surveyors were instructed to identify missing items and discontinuities or inconsistencies and to elaborate on the reasons for their absence in tag field 900, the abstract. Since breaks sometimes indicated significant landmarks in the PUD's history, records having been subpoenaed for instance, explanatory comments were very important. Interns were also instructed to give the location of missing records if known.

Function/Purpose/Comments-Tag Field 900

The abstracting field was used to record historical and functional information on agencies, offices and records series. In addition, this field was used to describe series inconsistencies, elaborate on information from tag fields 200, 203, 550 and 700, to indicate the location of duplicate series and the rate of reference if known. For agency cover sheets interns were instructed to include the date formed, date energized, number of customers, average annual sales and any comments on their history, organization and records situation which would be helpful in understanding the PUD: the oldest PUD in Oregon, the first generating PUD in Washington and so on. For an office or department, interns were to describe its function and duties, its place in the organization's structure (who were its subordinates and who its superiors), the title of the office head and any other comments which would illuminate its situation; for instance, when the office was created and why. For series

descriptions, interns entered the purpose or function of the record, its contents and archival value. Rates of reference were to be recorded, if known, and the existence of any finding aids was to be noted. Finally, interns were told to add any other information necessary for a complete and coherent description of the agency, office or series. The field length of nine lines of fifty-seven characters each was more than ample for our purposes.

A Note on Modification and Processing of Forms

The project used the OS-ALC version of the SPINDEX program. We introduced nine fields into the format with a total requirement of less than five hundred characters, well within the assigned limits. Data was entered with punch cards as this was the method our computer was arranged to receive. We believe that data entry with the IBM magnetic tape Selectric typewriter or similar brand-name devices offers advantages to cards. First, cards are cumbersome and create storage problems of their own. Second, they open the gate to high rates of error. Finally, they are more difficult to correct than other available means of data entry. We are moving to a more direct means of data entry in the second phase of our work.

In order that data could be checked before it was encoded on cards, we established procedures which would allow project office staff to complete much of the repetitive work. Field surveyors completed tag fields 110, 200, 201, 202, 203, 250, 300, 500, 550, 600, 601, 602, 700 and 900. Once they had completed an agency, interns checked their forms and returned them to the project office. Here the administrative staff again checked them for accuracy and then completed tag fields 000 (with the exception of Block 20) and 400. The forms were then sent to the University's data processing center. Punched cards were inspected and then entered on the project's master files. Corrections were made as necessary. With punch cards, corrections proved to be a rather combersome, though not particularly difficult, process.

On the whole, SPINDEX has worked relatively well for us.

RECORDS SERIES INVENTORY — **PACIFIC NORTHWEST PUBLIC POWER RECORDS SURVEY** — **FORM #1** 2/17/77 (Draft)

A. ORGANIZATION	B. OFFICE OF RECORD	C. CONTACT	D. TELEPHONE	E. LOCATION (ADDRESS) OF RECORD	F. NAME OF SURVEYOR	G. DATE

KEY-PUNCHED ITEMS

COMMON ENTRIES (Office Use Only)	1. RGN	5. OFF CODE	9. SERIES	13. DUMMY	19. R/A CODE	20. FA CODE
				0,0,0,0,0,0		

110. RECORD SERIES TITLE

200. CONTENTS

- ☐ AGREEMENTS
- ☐ AUDIO-TAPED
- ☐ BIDS
- ☐ CHARTS
- ☐ CONTRACTS
- ☐ DIRECTIVES
- ☐ DRAWINGS
- ☐ JOURNALS
- ☐ LEDGERS
- ☐ LETTERS
- ☐ MAPS
- ☐ OTHER (Specify)
- ☐ MEMORANDA
- ☐ MINUTES
- ☐ NEGATIVES
- ☐ ORDINANCES
- ☐ PHOTOGRAPHS
- ☐ PLANS
- ☐ PLATS
- ☐ REPORTS
- ☐ RESOLUTIONS
- ☐ VIDEO-TAPES

201. ARRANGEMENT

- ☐ ALPHABETICAL
- ☐ CHRONOLOGICAL
- ☐ GEOGRAPHICAL
- ☐ OTHER (Specify)
- ☐ NUMERICAL
- ☐ ORGANIZATIONAL
- ☐ SUBJECTIVE

202. RECORD-SIZE

- ☐ SMALLER_THAN_LETTER_SIZE
- ☐ LETTER_SIZE
- ☐ LEGAL_SIZE
- ☐ LARGER_THAN_LEGAL_SIZE

203. PHYSICAL-CHARACTERISTICS

- ☐ BOUND
- ☐ UNBOUND
- ☐ FLAT_FILED
- ☐ FOLDED
- ☐ ORIGINALS
- ☐ REPRODUCTIONS
- ☐ BOTH_ORIGINALS_AND_REPRODUCTIONS
- ☐ OTHER (Specify)

250. INCLUSIVE DATES — From / To

300. VOLUME OF RECORDS — Cubic Feet: C | Linear Feet: L | Microfilm Reels: M

400. LOCATION CODE

500. RECORD CLASSIFICATION

- ☐ OPR (Official Public Record)
- ☐ OFM (Office Files & Memoranda)

550. ARCHIVAL VALUE

- ☐ ARCHIVAL_VALUE
- ☐ NO_ARCHIVAL_VALUE

RECOMMENDED RETENTION

600. LEGAL	601. ORGANIZATION	602. SURVEYOR

700. SERIES INCONSISTENCIES

- ☐ FILE_BREAKS
- ☐ MISSING_ITEMS
- ☐ NO_FILE_BREAKS_OR_MISSING_ITEMS

FUNCTION/PURPOSE/COMMENTS

9,0,0	
9,0,1	
9,0,2	
9,0,3	
9,0,4	
9,0,5	
9,0,6	
9,0,7	
9,0,8	

-116-

Reduced from 8-1/2" x 14"

Appendix E

Brent G. Thompson, "Guide Project Tag Structure" [This paper was contributed to Session II but not presented verbally.]

The following is a list showing the tag structure used by the L.D.S. Church Historical Department in its basic finding aid. All the indexing tags have the function indicator 3 so they are field indexed as one term. The 100, 110, 400, 600, 900, 901, 902, 903, 904, and 905 tags are used for the Guide (ARSP5). The main index uses 110, 120, 200, and 210 tags as keywords with the 100, 110, 400, and 600 tags providing information under the keywords. The Special Lists index uses 700 and 730 tags for keywords with the 100, 110, and 710 tags appearing under the keywords.

Tag Number	Field Number	Description
000	Control Number	The control number is based on the collection number therefore MSD 1234 (which is a manuscript collection) has the control number A01234.
100	Title	The title includes the author or creator usually a generic title and span dates of the documents. An example is SMITH, JOSEPH, PAPERS, 1827-1844.
110	Significant Title	We use this tag only when a title needs to appear in the index.
120	Main Entry	This is the author or creator of the collection. Personal names are entered last name first. Birth and death dates are included when possible. An example is YOUNG, BRIGHAM, 1801-1877. The 120 tags never end with a period so that main entries and subject entries are separate in the index.
200	Supplied Subjects	This tag includes subjects from our authority list, place names, and buildings. (201-209 and 20A-20Z can also be used.)
210	Personal Name and Organization Subjects	Personal names can be entered as added entries (YOUNG, 1801-1877 - AUTHOR) or as subjects (YOUNG, BRIGHAM, 1807-1877). The names of Church auxiliaries and departments or corporate names would also

Tag Number	Field Number	Description
		be entered under this tag. Examples are DESERET SUNDAY SCHOOL UNION. and CANNON, GRANT AND COMPANY. (210-219 and 21A-21Z are used as needed.) The 210 tags always end in a period to separate them from a corresponding main entry.
400	Collection Number	The numbers are right justified so as to be more readable. Examples are MSDxxxxx72 and MSDxxx1234. (An x is used to indicate the number of spaces entered.)
600	Film Number	Many of the department's collections have been microfilmed. The film has a separate number. The fourth item on film 162 would be entered as Fxx162x#x4. (An x is used to indicate the number of spaces used.)
700	Donor	Personal names of donors are entered last name first. If the originals are in private possession or at another institution then one space proceeds the entry. Example is xYALE UNIVERSITY LIBRARY. The 700, 710, and 730 tags are used to produce a separate index we call "Special Lists."
710	Date Donated	Only the year is entered
730	Special Lists	Special Lists are produced by entering two spaces followed by a two digit abbreviation. Examples are: Journals--xxJR, Language--xxL:xxSPANISH, Restrictions--xxR:1995, Daguerreotypes--xxDG.
900	Collation	Volume, size, form. Examples are: xx1V. 37CM. MS., and xx3FT. MS.
901	First Content Note	This provides the researcher with brief information about the creator and a description of what material and potentially useful information the collection contains.

Tag Number	Field Number	Description
902	Second Content Note	Used if the content note needs to be divided into two paragraphs.
903	Donor/Source Note	The donor is listed if we have the originals. An example is xxGIFT OF JAMES D. MOYLE, SALT LAKE CITY, UTAH, 1975.
904	Restriction Notes	Explains any restrictions to use of the collection. Examples are xxCLOSED UNTIL 1995 EXCEPT BY PERMISSION OF THE DONOR, xxCLOSED EXCEPT BY PERMISSION OF THE CHURCH ARCHIVIST.
905	Finding Aids Note	An example is: xxSEE REGISTER FOR COMPLETE LIST.

Appendix F: South Carolina Department of Archives and History Exhibits
General Assembly
Petitions

000 0010 003 Year NDOC Number 00380 Item 00

100 McCREIGHT, WILLIAM COMMANDER OF THE WINNSBORO LIGHT INFANTRY VOLUNTEERS, PETITION FOR THE PASSAGE OF A LAW GRANTING TO UNIFORM COMPANIES THE PRIVILEGE OF RECRUITING MEN FROM ANY REGIMENT WITHIN THEIR RESPECTIVE BRIGADES.

(4 pages)

250 Date NODO

Topics

200 WINNSBORO LIGHT INFANTRY VOLUNTEERS
201 MILITIA AND PATROL LAWS
202
203
204
205

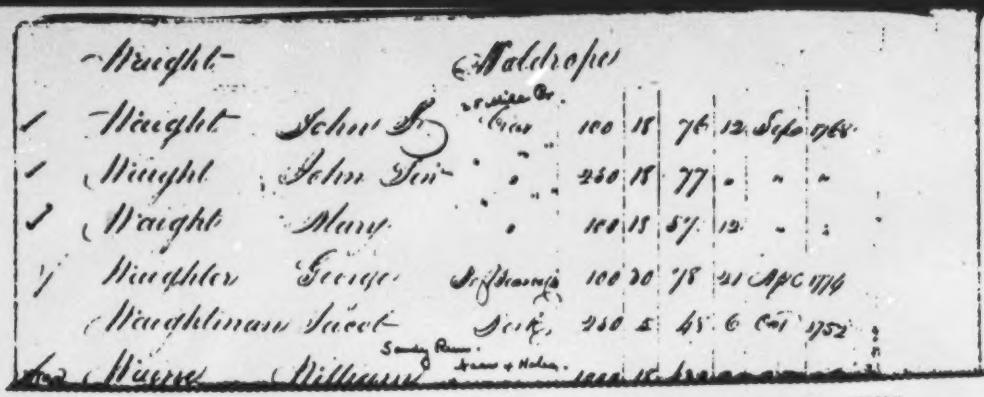
Locations

500 FAIRFIELD DISTRICT
501
502
503
504
505
506
507

Personal Name Titles

120 McCREIGHT, WILLIAM
121 PEAYS, COL.
122 DAVIS, COL.
123
124
125
126
127
128
129
210
211
212
213
214
215
216
217
218
219

Example # 1



Wright — Waldrop

✓	Wright	John Sr.	Craven Co.	100	18	76	12	Sept	1768
✓	Wright	John Jun.	" "	250	18	77	"	"	"
✓	Wright	Mary	"	100	18	87	12	"	"
✓	Waughter	George	Northampton	100	20	78	21	Apr	1776
	Waughtman	Jacob	Berk.	250	£	45	6	Oct	1752
✓	Waine	William	Craven Co.	1000	18				

← Original index

Index converted into Spindex format on an edit listing

05/31/77

PAGE 136

SOUTH CAROLINA DEPARTMENT OF ARCHIVES AND HISTORY

INDEX INPUT LIST

REC GRP	SERIES	BX/VOL	FOLD/PG	ITEM	ACTION	TYPE	TAG	DATA ELEMENT
3332	335	0013	00077	00	3	A	000	
3332	335	3018	00077	33	1	A	100	WAIGHT, JOHN SR., LAND GRANT FOR 250 ACRES IN CRAVEN COUN
0032	335	0014	00077	00	9	1	100	TY.
3332	335	3019	00077	00	1	A	250	1768/09/12
3332	335	3018	00077	00	1	A	600	000250 ACRES
3032	335	0014	00077	00	1	A	500	CRAVEN COUNTY
3332	335	3018	00077	33	1	A	120	WAIGHT, JOHN SR.
3032	335	3019	00057	00	3	A	000	
3002	335	3018	00057	33	1	A	100	WAIGHT, MARY, LAND GRANT FOR 100 ACRES IN CRAVEN COUNTY.
3332	335	3019	00057	35	1	A	250	1769/09/12
3332	335	3016	00057	33	1	A	600	000100 ACRES
3002	335	3014	00057	33	1	A	500	CRAVEN COUNTY
3332	335	3019	00057	00	1	A	120	WAIGHT, MARY
3002	005	0035	00074	00	3	A	000	
3332	335	3017	00374	00	1	A	100	WAIGHTEN, GEORGE, LAND GRANT FOR 100 ACRES ON BEAVER CREE
3332	335	3333	00374	33	9	1	100	K.
3332	335	0033	00374	00	1	A	250	1774/04/21
3332	335	3333	00374	33	1	A	600	000100 ACRES
3332	335	3333	00374	33	1	A	500	BEAVER CREEK
3332	335	3333	00374	33	1	A	120	WAIGHTEN, GEORGE
3332	335	3035	00364	00	3	A	000	
3332	335	3035	00364	33	1	A	100	WAIGHTMAN, JACOB, LAND GRANT FOR 250 ACRES IN BERKELEY COU
3032	335	0015	00364	33	9	1	100	NTY.
3332	335	3035	00364	33	1	A	250	1752/10/06
3332	335	3035	00364	33	1	A	600	000250 ACRES
3332	335	3035	00364	33	1	A	500	BERKELEY COUNTY
3332	335	3035	00364	00	1	A	120	WAIGHTMAN, JACOB
3332	335	3014	00437	00	3	A	000	
3332	335	3014	00437	33	1	A	100	WAIN., WILLIAM, LAND GRANT FOR 1,000 ACRES IN BERKELEY COU
3032	335	0015	00437	33	9	1	100	NTY.
3332	335	3014	00437	00	1	A	250	1769/03/25
3332	335	3014	00437	33	1	A	600	001000 ACRES
3332	335	3014	00437	33	1	A	500	BERKELEY COUNTY
3332	335	3014	00437	33	1	A	120	WAINE, WILLIAM
3332	335	3014	00433	33	3	A	000	
3332	335	3014	00433	33	1	A	100	WAINWRIGHT, SAMUEL, LAND GRANT FOR 43 ACRES IN ST. PAULS P
3032	335	0015	00433	33	9	1	100	ARISH.
3332	335	3014	00433	00	1	A	250	1769/07/10
3332	335	3014	00433	33	1	A	600	000043 ACRES
3332	335	3014	00433	33	1	A	500	SAINT PAULS PARISH
3332	335	3014	00433	33	1	A	120	WAINWRIGHT, SAMUEL
3332	335	3035	00367	33	3	A	000	
3332	335	3035	00367	33	1	A	100	WAINWRIGHT, SAMUEL, LAND GRANT FOR 600 ACRES IN CRAVEN COU
3032	335	0015	00367	33	9	1	100	NTY.

-121-

134

135

Example #2

Appendix G

Joel A. Shirley, "Designing Tag Structures" [This paper was contributed to Session II but not presented verbally.]

The South Carolina Department of Archives and History is using Spindex for two separate applications. For each of these applications we have developed our own tag structure and usage.

The tag structure in our document index is as follows:

100 tag--Description

This tag is a fixed format, alphanumeric description of the document. Essentially a very brief abstract of the document that we print out along with each index reference, this tag includes in its description the primary personal name(s), type of document, main topic(s), and geographic location. This tag is limited to 570 positions (letters, numbers, punctuation and spaces).

110 tag--Type of Document

This tag contains the legal type of document (e.g. plat, lease and release, petition, etc.). We use Words and Phrases and Black's Law Dictionary as our basic guides. Type of document is limited to 50 positions.

120-129, 12A-12Z,
210-219, 21A-21Z tags--Personal Names

As can be guessed, we merged the 120 and 210 tags and in our tag structure do not distinguish between principal parties involved in a document and persons mentioned. Since we print out with each index reference a brief description of the document that includes the names of the principal parties, this poses no problem for the user. We use the 120 and 210 tags for all personal names, surname first. There is a possibility of 72 personal names per indexed item, each having 50 positions.

200-209, 20A-20A tags--Main Topics

We reserve these tags for topical entries, selected from our subject thesaurus. Although we seldom use more than a few topics for a given item in the document index, we have allowed for the inclusion of 36 topics, each with 50 positions.

250 tag--Date

This tag is maintained as the 21 position date tag. We use exact dates, span dates, circa dates, and no dates (N.D.).

500-509, 50A-50Z tags--Locations

The 500 tag is used for geopolitical and geophysical location of the main topic of the document instead of using the 200 tag. By using an assigned tag we are able to be more selective of information entered into the other tags. There is a possibility of 36 locations each having 50 positions.

600-602 tags--Quantity

Although we do not use this tag for output at present, we do enter information from a conveyance (grant, deed, bill of sale, etc.) concerning the number of acres, town lots, or slaves being transferred. This information will be used in the future for data retrieval.

The tag structure in our prose descriptive guide to record series, subgroups and record groups is as follows:

100 tag--Title

This is reserved for the record group, subgroup, or series title. The tag is limited to 100 positions or two lines of print, 50 characters each.

200-209, 20A-20Z tags--Topics

In our descriptive guide these tags contain any type of topical entry including personal names, geographic locations and type documents. There are 36 possible topics, each having 50 positions.

250 tag--Date

This tag is limited to 9 positions and indicates the year date or dates of the existing records contained in the record group, subgroup, or series. Most entries are span dates (i.e. 1827-1915).

300 tag--Quantity

This tag contains the physical quantity of the record group (in subgroups or series), the subgroup (in series), or the series (in abbreviations for type or types of storage containers, such as volumes, flat files, record center cartons, audio tapes, etc.). This tag is limited to 25 positions.

700 tag--Entry Number

This tag has been reserved for a 14 position entry number which reflects the repository ID, section ID (the large subdivision of the holdings), and archival arrangement of the records series. We chose to add a tag instead of using the 000 (Control Field) because of possible rearrangement of records. A tag correction only requires one computer step and it will therefore be easier and cheaper to rearrange the entries in our descriptive guide by using this tag to arrange the entries.

900, 910, 920, 930, and 940 tags--Description

These tags are limited to 570 characters each and are the five paragraphs of description that are the heart of the descriptive guide. Basically these paragraphs contain a full archival description of the record group, subgroup, or series.

For both of our applications we use the same 000 tag structure. Instead of the standard Spindex II control number subfield allocations of 4, 4, 4, 3, 3, we changed them to 4, 3, 4, 5, 2. They are Record Group (4 positions); Series (3 positions); Box, volume, year, or Microfilm Roll Number (4 positions); Folder, Page, Document Number, or Frame Number (5 positions); and Item Number (2 positions).

SPINDEX INPUT
GENERAL PURPOSE

Control Field
000 LVL II _____ LVL III ___ LVL IV _____ LVL V _____ LVL VI __

Data Element
100 _____

_____

_____

_____

_____

_____

250 Date

600 Quantity: _____

Topics

200 _____

201 _____

202 _____

Type Document

110 _____

Locations

500 _____

501 _____

502 _____

503 _____

504 _____

505 _____

506 _____

507 _____

508 _____

509 _____

Personal Name Titles

120 _____

121 _____

122 _____

123 _____

124 _____

125 _____

126 _____

127 _____

128 _____

129 _____

210 _____

211 _____

212 _____

213 _____

214 _____

215 _____

216 _____

217 _____

218 _____

219 _____

Initials _____ Date _____

Descriptive Guide Input

Control Field
3A 000 LVL II _ _ _ _ LVL III _ _ _ LVL IV 0 0 0 0 LVL V 0 0 0 0 0 LVL VI 0 0

TITLE
1A 100 _____

_____

TOPICS

1A 200 _____ 1A 205 _____

1A 201 _____ 1A 206 _____

1A 202 _____ 1A 207 _____

1A 203 _____ 1A 208 _____

1A 204 _____ 1A 209 _____

DATE

1A 250 _____

QUANTITY

1A 300 _____

ENTRY NUMBER

1A 700 SCA/1 _ _ _ _ _ _ _ _

DESCRIPTION

1A 900 ~~333333~~ _____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

INITIALS_____ DATE_____